SF 298 MASTER COPY KEEP THIS COPY FOR REPRODUCTION PURPOSES

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB NO. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comment regarding this burden estimates or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

| 1. AGENCY USE ONLY *(Leave blank)* | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | April 16, 1998 | Final report: Sep. 1, 1995 - Dec 31, 1997 |

**4. TITLE AND SUBTITLE**
Artificial Quantum Solids: Physics, Fabrication and Applications

**5. FUNDING NUMBERS**
DAAH04-95-1-0586

**6. AUTHOR(S)**
S. Bandyopadhyay and A. E. Miller

**7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)**
Univ. of Notre Dame
Notre Dame, IN 46556

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**
ARO 34619.18-PH

**11. SUPPLEMENTARY NOTES**
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

**12a. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for public release; distribution unlimited.

**12 b. DISTRIBUTION CODE**

**13. ABSTRACT** *(Maximum 200 words)*

This report describes research carried out at the University of Nebraska and at the University of Notre Dame in the physics and fabrication of self-assembled arrays of quantum dots and wires. Such arrays have shown enhanced optical non-linearities, novel giant magnetoresistance and other intriguing properties.

DTIC QUALITY INSPECTED 2

19980519 124

**14. SUBJECT TERMS**
Quantum dots, quantum wires, self-assembly, non-linear optics, magnetotransport, quantum computing networks.

**15. NUMBER IF PAGES**
6+appendix

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OR REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| UNCLASSIFIED | UNCLASSIFIED | UNCLASSIFIED | UL |

# Artificial Quantum Solids[1]

## FINAL REPORT

S. Bandyopadhyay and A. E. Miller

April 1998

US Army Research Office

CONTRACT NO. DAAH04-95-1-0586

APPROVED FOR PUBLIC RELEASE
DISTRIBUTION UNLIMITED

THE VIEWS, OPINIONS, AND/OR FINDINGS CONTAINED IN THIS REPORT ARE THOSE OF THE AUTHORS AND SHOULD NOT BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, POLICY, OR DECISION, UNLESS SO DESIGNATED BY OTHER DOCUMENTATION.

[1]This contract was awarded to University of Notre Dame, the PI's previous employer. The PI is presently with the University of Nebraska and can be reached by e-mail at the address bandy@engrs.unl.edu or by telephone at (402) 472-0294. His mailing address is Department of Electrical Engineering, University of Nebraska, Lincoln, Nebraska 68588-0511.

FINAL REPORT
presented to the
US ARMY RESEARCH OFFICE
on contract DAAH04-95-1-0586
Artificial Quantum Solids:
Physics, Fabrication and Applications

# 1 Foreward

This report covers the period September 1, 1995 to December 31, 1997. This was the duration of grant DAAH04-95-1-0586. During this time, major advancements were made in the development of self-assembled nanostructures. We developed two electrochemical techniques for self-assembling *periodic* arrays of quantum dots. These arrrays are (at the time of writing this report) the most ordered self-assembled arrays reported. They have been applied to fabricate non-linear optical elements with vastly enhanced second-order susceptibility. Additonally, we have developed quantum neuromorphic network paradigms for collective computation that are based on charge inetractions in such arrays.

Two graduate students were supported under this program at the University of Notre Dame. One received a Ph.D. degree in Electrical Engineering in December 1996 and is currently a post-doctoral research associate at the University of California-Los Angeles. The other is still awaiting completion of his dissertation.

This research resulted in about fifteen publications and numerous talks and colloquia.

A Table of Contents is not included since the main body of the report is less than 10 pages.

# 2 Technical Report

## 2.1 Statement of problem

The research sponsored by this grant addressed the electrochemical self-assembly of quantum wires and dots. The motivation was to develop a process that leads to *ordered* and periodic arrays as opposed to random arrays with large size dispersion. Periodic arrays can have interesting applications in transport, computing networks, non-linear optical devices and photonic crystals.

## 2.2 Summary of most important results

The following findings were the most remarkable:

1. We found that electropolishing an aluminum foil in a solution of ethanol (70.0 vol. %), distilled water (13.8 vol. %), butyl cellusolve (10 vol. %) and perchloric acid (6.2 vol %) can produce a regimented array of crests and troughs on the surgace of the aluminum foil. If the electropolishing is carried out at 60 volts for 30 seconds, then an array of crests and troughs form that can be used to produce quantum dots. If the electropolishing is carried out at 50 volts for 10 seconds, an array of stripes form that can be used to produce quantum wires. Raw atomic force microgrpahs of these arrays are shown in the next section.

2. When an electropolished aluminum surface is anodized in sulfuric acid under a current density of 40 $mA/cm^2$, a nanoporous alumina film forms on the surface with a quasi-periodic array of pores with an average diameter of 100 Å. These pores can be filled up with the material of interest by ac electrolysis, thereby forming a quantum dot array dispersed in alumina. These dots have been directly imaged by TEM, SEM, cross-section TEM. field-emission SEM, and characterized by Auger, ESCA, energy-dispersive analysis of x-ray, Raman spectroscopy, photoluminescence, magnetotransport measurements, susceptometry, ellipsometry and non-linear pump-and-probe spectroscopy. These measurements have revealed strong signatures of quantum confinement.

3. We have observed novel giant magnetoresistance in nickel quantum dot arrays caused by remote spin-flip scattering

4. We have observed a five-fold increase in the second-order non-linear suscpetibiltiy of CdS quantum dots over bulk CdS

5. We have observed photoluminecence from CdS quantum dots with an efefctive optical diameter of 35 Å. These are some of the smallest semiconductor dots to show optical activity.

6. We have designed computing architectures based on these structures.

Much of this research was carried out in collaboration with Dr. R. E. Ricker of the National Institute of Standards and Technology, Gaithersburg, Dr. Jeffrey A. Eastman of Argonne National Laboratory, Professor Meera Chandrasekhar of the University of Missouri-Columbia, Dr. David Janes of Purdue University, Prof. Vwani Roychowdhury of the University of California Los Angeles, Prof. Paul Snyder of the University of Nebraska and Prof. H-C Chang of the University of Notre Dame.

## 2.3 Personnel

1. Prof. Supriyo Bandyopadhyay, principal investigator
2. Prof. Albert E. Miller, co-principal investigator
3. Dr. Alexander Balandin, currently post-doctoral research associate at the University of California-Los Angeles
4. Mr. Vadim Yuzhakov, Ph.D. student in the Department of Chemical Engineering, University of Notre Dame, IN

Dr. Balandin was a graduate student working on this grant who received his Ph.D. in December 1996.

# APPENDIX

Publications.

*Superlattices and Microstructures, Vol. 23, No. 3, 1998*

# Magnetostatic modulation of nonlinear refractive index and absorption in quantum wires

A. BALANDIN, S. BANDYOPADHYAY†
*Department of Electrical Engineering, University of Nebraska, Lincoln, NE 68588, U.S.A.*

*(Received 15 July 1996)*

The magnetic-field dependence of the nonlinear differential refractive index $\Delta n$ and absorption $\Delta\alpha$ in quantum wires—measured by non-degenerate pump and probe spectroscopy—is investigated theoretically. The nonlinearities arise from population saturation of the excitonic state under optical pumping and the formation of biexcitons (excitonic molecules). Both $\Delta n$ and $\Delta\alpha$ exhibit positive and negative peaks at certain pump and probe detuning frequencies associated with the formation of biexcitons and bleaching of excitons, respectively. The amplitude, lineshape and the frequency at which these peaks occur can be modulated by a magnetic field which opens up the possibility of realizing novel magneto-optical devices. Additionally, the magnetic field may allow us to realize a relatively large variation in the differential refractive index over a range of frequencies without significant accompanying absorption, thereby allowing the observation of optical bistability.

© 1998 Academic Press Limited

Please supply [Key words:

## 1. Introduction

It is well known that quantum-confined structures exhibit pronounced optical nonlinearities of excitonic origin [1]. The enhanced nonlinearities arise from one-dimensional quantum confinement which increases the binding energy of all excitonic complexes and the oscillator strengths for excitonic transitions. In this paper, we report how a magnetic field influences the nonlinear differential refractive index $\Delta n$ and absorption $\Delta\alpha$ in a quantum wire. This study is motivated by the realization that any significant modulation of these quantities by a magnetic field can lead to novel device applications, as well as provide a tool for probing the origin of optical nonlinearity in a quantum structure.

The physical processes associated with non-linear refraction and absorption in quantum confined systems is a well-researched topic. In reference [2, 3], the authors reported room-temperature measurements of $\Delta n$ in GaAs multiquantum well structures and found it to be 0.01 at low levels of excitation and 0.05 at high levels. They attributed the non-linearity to band filling. In quantum wires (as opposed to wells), we can expect $\Delta n$ and $\Delta\alpha$ to be much larger because of the additional degree of confinement and the much higher density of states at the subband edges. Indeed, our theoretical calculations indicate that $\Delta n$ can be an order of magnitude larger in quantum wires than that found in quantum wells.

† The authors are on leave from the University of Notre Dame, Notre Dame, IN 46556, U.S.A.

0749–6036/97/030001 + 07 $25.00/0 sm960258

© 1998 Academic Press Limited

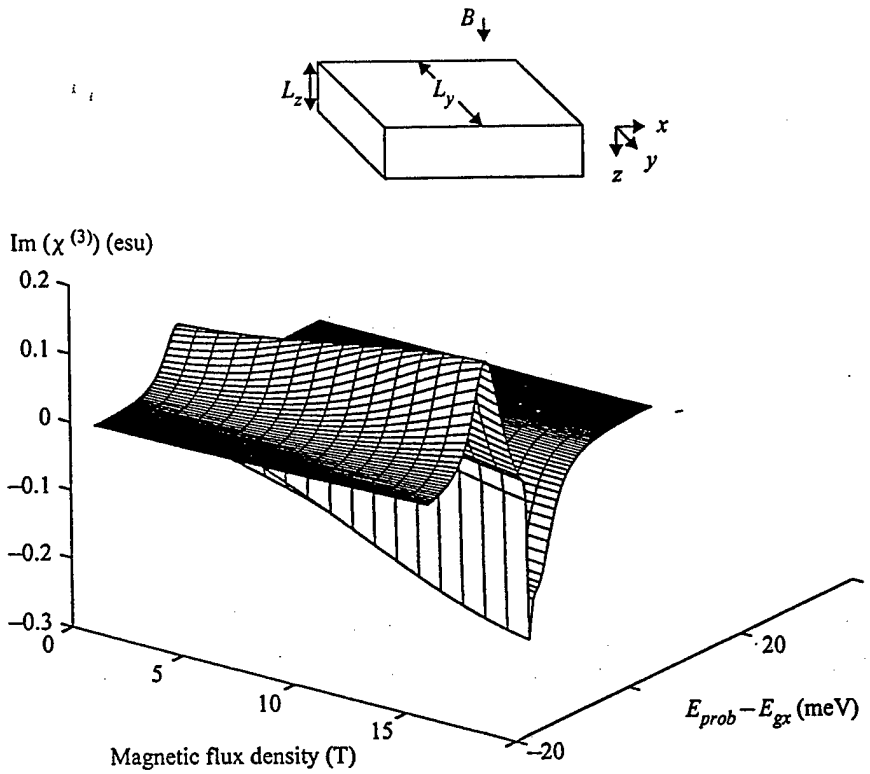



**Fig. 1.** The imaginary part of the third-order nonlinear susceptibility $\chi^{(3)}$ as a function of pump and probe detuning energy and magnetic-flux density. The pump is tuned slightly below the exciton resonance of each value of the magnetic field and the longitudinal broadening parameter is assumed to be one-tenth that of the transverse broadening parameter. The wire dimension is $L_z = 200$ Å and $L_y = 500$ Å. The inset shows the wire geometry.

## 2. Theory

We wish to calculate the differential refractive index and absorption associated with the third-order nonlinear susceptibility $\chi^{(3)}$ in quantum wires. For this, we consider a rectangular quantum wire of the geometry shown in the inset of Fig. 1. An external magnetic field is applied perpendicular to the wire axis. We assume near-resonant pumping of the excitonic state in a non-degenerate pump and probe spectroscopy experiment and calculate the changes in refractive index $\Delta n$ and absorption $\Delta\alpha$ relevant to this situation. The actual measurable quantities in such an experiment are usually the transmission in the absence ($T_0$) and in the presence ($T$) of the pump. The differential transmission spectra can be found from these quantities as $D = (T - T_0)/T_0$. For small values of the differential transmission (well below unity), $D$ is proportional to the differential absorption $\Delta\alpha$. In fact, $D \approx -\Delta\alpha d$, where $d$ is the wire thickness along the direction of the optical beam.

The non-linear differential refractive index and absorption can be evaluated theoretically as in [4]. These quantities are given by

$$\Delta n = \frac{2\pi}{\sqrt{\epsilon_r}} \mathrm{Re}\chi^{(3)}, \tag{1}$$

and

$$\Delta\alpha = \frac{4\pi\omega}{c\sqrt{\epsilon}} \mathrm{Im}\chi^{(3)}, \tag{2}$$

**Table 1:** Values of the various parameters for GaAs used to calculate the non-linear susceptibility $\chi^{(3)}$.

| |
|---|
| $E_{g0} = 1.519$ eV |
| $\hbar\Gamma = 3$ meV |
| $E_p = 23$ eV |
| $N_0 = 7.89 \times 10^{14}$ cm$^{-2}$ |

where $c$ is the speed of light, $\epsilon_r$ is a relative dielectric constant of the material, $\omega$ is a near-resonant frequency of the pump beam, and $\mathrm{Im}\chi^{(3)}$, $\mathrm{Re}\chi^{(3)}$ are the imaginary and real parts of the non-linear third-order susceptibility $\chi^{(3)}$ which need to be calculated.

The general derivation of $\chi^{(3)}$ for low density of excitonic complexes can be found in [5]. This derivation is based on summation over 16 double Feynman diagrams. In the frequency range of interest, the lowest-lying states are the major contributors to $\chi^{(3)}$ and this allows us to reduced the expression for $\chi^{(3)}$ to a simplified form given by

$$\chi^{(3)} = \frac{-2}{\pi\sqrt{2\pi}}\frac{\tau}{\eta^2}\frac{N_0 e^4}{m_0^2\omega_{g_0}^4}E_p^2\left[\frac{1}{(\omega_1-\omega_{g_0}+i\Gamma_{g_0})}-\frac{1}{(\omega_1-\omega_{g_0}+\omega_b+i\Gamma_{bg})}\right]$$
$$\times\sum_{r=1}^{2}\left\{\frac{1}{\hbar^3(\omega_r-\omega_2+i\gamma)}\left[\frac{1}{(\omega_{g_0}-\omega_2+i\Gamma_{g_0})}+\frac{1}{(\omega_r-\omega_{g_0}+i\Gamma_{g_0})}\right]\right\}$$
$$+\frac{1}{(\omega_1+\omega_2-2\omega_{g_0}+\omega_b+i\Gamma_{b_0})}\left[\frac{1}{(\omega_1-\omega_{g_0}+i\Gamma_{g_0})}+\frac{1}{(\omega_2-\omega_{g_0}+i\Gamma_{g_0})}\right], \qquad (3)$$

where $\omega_2$ and $\omega_1$ are the pump and probe frequencies, $\hbar\omega_{g_0}$ is the exciton ground-state energy, $\hbar\omega_b$ is the biexciton binding energy, $m_0$ is the rest mass of a free electron, and $N_0$ is the average areal density of unit cells. The quantities $\Gamma_{ij}$ and $\gamma$ are the transverse and longitudinal broadening parameters (or damping constants), and $E_p$ is the Kane matrix element. The indices $i$ or $j$ indicates system ground state (0), exciton ground state ($g$), and biexciton ground state ($b$). Numerical values of the various quantities used in our calculations are given in Table 1. Parameters $\eta$ and $\tau$ physically correspond to the exciton and biexciton correlation lengths (electron–hole and hole–hole mean separations in the two cases) and have to be determined variationally for each magnetic field strength and for each set of wire dimensions following the prescription given in [6, 7].

The exciton ground-state energy $\hbar\omega_{g_0}$ is defined as follows

$$E_g^X = \hbar\omega_{g_0} = E_G + E_{e_1} + E_{hh_1} - E_B^X, \qquad (4)$$

where $E_G$ is a bulk band gap of the material, $E_{e_1}$, $E_{hh_1}$ are the lowest electron and the highest heavy-hole magnetoelectric subband bottom energies in a quantum wire (measured from the bottom of the bulk conduction band and the top of the bulk valence band) respectively, and $E_B^X$ is the ground-state exciton binding energy which is also determined variationally [6, 7].

It should be noted from eqn (3) that $\chi^{(3)}$ is a strong function of the transverse and longitudinal broadening parameters $\Gamma_{ij}$ and $\gamma$. Physically, $\gamma$ is related to the population decay rate of the excitonic states. The smaller the value of $\gamma$, the larger the lifetime of excitons and the higher the probability of forming a biexciton in a two-step photon absorption. The transverse broadening parameters $\Gamma_{ij}$ represent, for $i \neq j$, the phenomenological coherence decay rate of the $i-j$ transition, while for $i = j$, they describe the population decay of the state $i$. The population decay rate, in its turn, is determined by the dominant scattering mechanism in the sample. In most cases, the values of $\Gamma_{ij}$ and $\gamma$ are difficult to obtain experimentally and fairly difficult to estimate theoretically. Moreover, these parameters could be strong functions of the confinement, population density of

excitons, magnetic field and temperature. In view of little experimental data available, and in order to simplify the calculations, we assume that $\Gamma_{ij} = \Gamma$ for all $i, j$.

Since in this work we are interested in the modulation of the differential refractive index and absorption of quantum wires with a magnetic field, the influence of the field on all parameters in eqn (3) is especially important. The value of $\Gamma$ in quantum wires is primarily determined by carrier–phonon interactions [8]. As shown in [8], the scattering rates associated with these interactions can be affected by a magnetic field at any given kinetic energy of an electron or hole. However, when the rates are averaged over energy, the magnetic-field dependence turns out to be quite weak. As a first approximation, we can therefore consider the rates to be independent of the magnetic field. We also neglect thermal broadening of the damping parameters since it is less important in quantum-confined systems than in bulk [9]. An important property of equation (3) is the following: if all the transverse relaxation parameters are assumed to be equal (as in our case) and the biexciton binding energy ($\hbar\omega_b$) approaches zero, then $\chi^{(3)}$ vanishes. This is a manifestation of the well-known fact that non-interacting ideal independent bosons do not show any non-linearity [9]. Consequently, exciton–exciton interaction, leading to biexciton formation, is necessary for the existence of this type of the non-linearity.

A calculation of the excitonic contribution to $\chi^{(3)}$ requires that the exciton and biexciton binding energies be obtained first. Additionally, all the parameters $\eta$ and $\tau$ need to be found. For details of computing these energies and these parameters in the case of a quantum wire subjected to a magnetic field, we refer the reader to our past work [6, 7]. Once these quantities are evaluated, we can calculate $\chi^{(3)}$ from equation (3) as a function of a magnetic field, wire width and pump and probe detuning frequencies. The differential refractive index and absorption are then computed from the real and imaginary parts of $\chi^{(3)}$ as given by eqns (1) and (2).

## 3. Results and discussion

All results in this paper are pertinent to GaAs quantum wires. In Fig. 1, we present a three-dimensional plot of $\mathrm{Im}\chi^{(3)}$ for a two-beam experiment in which the frequency of one beam, the pump, is fixed and that of the other, the probe, is allowed to vary over a frequency range of $\hbar\Delta\omega = 40$ meV centered around the pump frequency. The pump frequency is chosen to be slightly detuned from the exciton resonance by a frequency $-\frac{\sqrt{2}}{2}\Gamma/\hbar$. The quantum-wire dimensions which have been used to plot this figure are $L_y = 500$ Å, $L_z = 200$ Å. The longitudinal broadening parameter $\gamma$ is chosen to be one-tenth that of the transverse broadening parameter $\Gamma$ which is a physically reasonable ratio.

A pronounced negative peak is present in the spectrum for all values of a magnetic field. It represents strong transmission which is due to a saturation (or bleaching) of the excitonic state. Physically, the initial exciton population created by the pump beam tends to amplify the probe beam when its energy is tuned at or near the exciton ground state (this corresponds to the linear gain peak). A magnetic field makes the peak deeper, without significant broadening, thus enhancing transmission further. Another feature of interest is in the region of positive $\mathrm{Im}\chi^{(3)}$ that corresponds to optical *absorption*. This absorption may be attributed to the formation of an excitonic molecule (biexciton). The initial exciton population enables the probe to be more strongly absorbed when its energy matches the exciton–biexciton transition energy $\hbar(\omega_{g_0} - \omega_b)$.

The same basic features are repeated in the absorption spectrum presented in Fig. 2. Here we plot the differential absorption $\Delta\alpha$ as a function of the pump and probe detuning frequencies when the longitudinal broadening parameter $\gamma$ is one-tenth of the transverse broadening parameter $\Gamma$. As we can see, when the pump frequency is nearly resonant with the excitonic absorption, the swing in the differential absorption $\Delta\alpha$ is very large ($0.5 \times 10^5$ cm$^{-1}$–$10^5$ cm$^{-1}$). Another feature to note is that the frequency separation between the positive and negative peaks (associated with biexciton formation and exciton bleaching) is quite sensitive to the magnetic field. This separation is not sensitive to damping (values of $\gamma$ and $\Gamma$) or slight detuning of the pump. Therefore, we can use a magnetic field to tune this separation, thus realizing magneto-optical devices.

In Fig. 3, we show the differential refractive index $\Delta n$ as a function of the pump and probe detuning

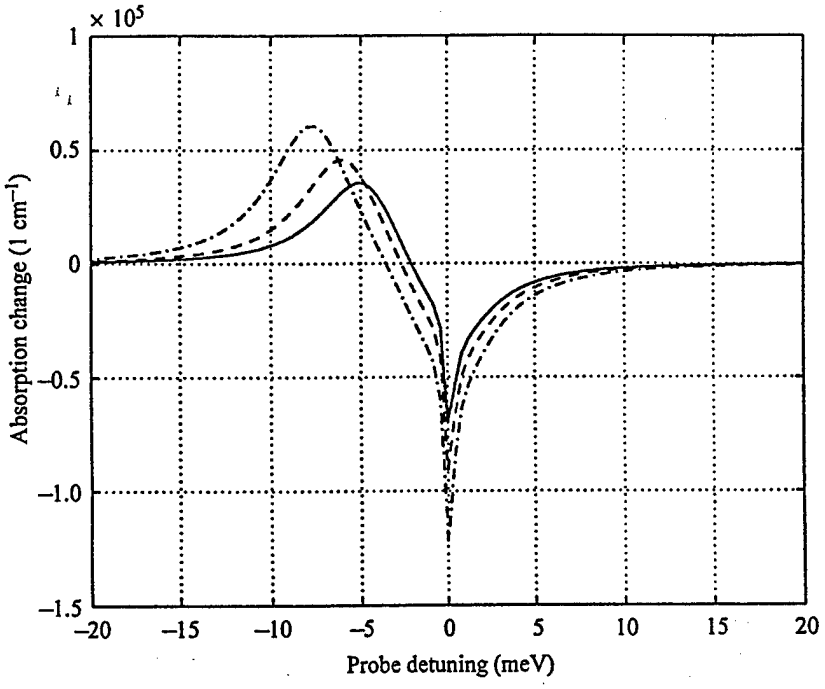



**Fig. 2.** The differential absorption $\Delta\alpha$ as a function of pump and probe detuning energies for different values of a magnetic field. The pump is set at exciton resonance for each value of a magnetic field. The longitudinal broadening parameter is one-tenth that of the transverse broadening parameter.

frequency and the magnetic field. More complicated behaviour is exhibited by $\Delta n$, with a strong negative peak occurring at the energy between the positive and negative resonances in the absorption change. The negative peak is related to the fact that $\Delta\alpha$ has a positive dispersive peak on its low-energy side.

Although not shown in this paper, we also found that damping has a deleterious effect on the nonlinearity. As the damping parameter $\gamma$ increases from $0.1\Gamma$ to $\Gamma$, the swing in $\Delta n$ drops from 0.4 to 0.05 when no magnetic field is present, resulting in a 20-fold reduction in the nonlinearity. However, when a magnetic flux density of 10 T is present, $\Delta n$ drops by only a factor of 6. Therefore, a magnetic field makes the non-linearity less sensitive to damping.

The strong dependence of $\Delta n$ and $\Delta\alpha$ on an external magnetic field has an important consequence for device applications. One possible application of band-gap resonant optical nonlinearities in quantum-confined systems is *optical bistability* and switching devices associated with it. Miller *et al.* [2] pointed out that in order to achieve optical bistability, one should provide a *large refractive index swing* at a relatively *low absorption* level. For bistable etalons using quantum wells, the relationship between minimum index change and absorption in the material for bistability to be observable can be written as $\Delta n/\alpha\gamma > \sqrt{3}/6\pi$, where $\lambda$ is the wavelength of the pump beam. Using this criterion, Miller *et al.* [2] concluded that bistability is not achievable in quantum-well etalons from excitonic mechanisms alone since in the region of large $\Delta n$, excitonic absorption is also very high. However, in quantum wires, the criterion for bistability can be met, especially in the presence of a magnetic field. This is a significant advantage.

## 4. Conclusion

In conclusion, we have investigated the dependence of $\Delta n$ and $\Delta\alpha$ in a quantum wire on an external

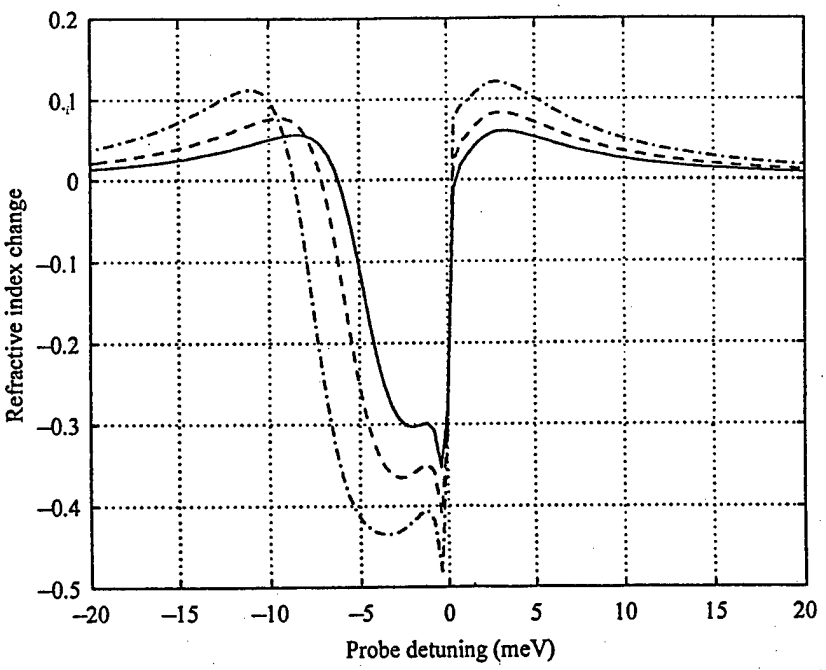



**Fig. 3.** The differential refractive index $\Delta n$ as a function of pump and probe detuning energies for different values of a magnetic field. All parameters and conditions are the same as in Fig. 2.

magnetic field. We found that the field makes these differential parameters less sensitive to damping and may make it possible to observe optical bistability. Additionally, the field can modulate the spectral characteristics of $\Delta n$ and $\Delta\alpha$ which may have device applications.

*Acknowledgements*—The authors are indebted to Dr Madarasz, University of Alabama, Huntsville and Dr Dneprovskii, Moscow State University for many insightful discussions. AB acknowledges support of the University of Notre Dame Zahm Research Travel Fund. This work was partially supported by the Army Research Office under Grant number DAAH04-95-1-0586.

## References

[1] See, for example, D. S. Chemla, D. A. B. Miller, and S. Schmitt-Rink, in *Optical Nonlinearities and Instabilities in Semiconductors*, edited by Hartmut Haug (Academic Press, New York, 1988) and references therein; D. S. Chemla, D. A. B. Miller, and P. W. Smith, in *Semiconductors and Semimetals*, edited by Raymond Dingle (Academic Press, San Diego, 1987), Vol. 24, Chap. 5.

[2] A. Miller, P. K. Milsom, and R. J. Manning in *Nonlinear Optics and Optical Computing* edited by S. Martellucci, and A. N. Chester (Plenum Press, New York, 1990) p. 119.

[3] N. Peyghambarian, S. W. Koch, H. M. Gibbs, and H. Haug in *Nonlinear Optics and Optical Computing* edited by S. Martellucci, and A. N. Chester (Plenum Press, New York, 1990) p. 99.

[4] L. Banyai, Y. Z. Hu, M. Lindberg, and S. Koch, Phys. Rev. **B38**, 8142 (1988).

[5] F. L. Madarasz, F. Szmulowicz, F. K. Hopkins, and D. L. Dorsey, Phys. Rev. **B49**, 13528 (1994); F. L. Madarasz, F. Szmulowicz, F. K. Hopkins, and D. L. Dorsey, Phys. Rev. **B52**, 8964 (1995); F. L. Madarasz, F. Szmulowicz, F. K. Hopkins, and D. L. Dorsey, J. Appl. Phys. **75**, 639 (1994).

Please update

[6] A. Balandin and S. Bandyopadhyay, Phys. Rev. **B52**, 8312 (1995); A. Balandin and S. Bandyopadhyay, Phys. Rev. B (in press).
[7] A. Balandin and S. Bandyopadhyay, Superlat. Microstruct. **19**, 97 (1996). For the influence of a magnetic field on electron and hole wavefunctions see also A. Balandin, and S. Bandyopadhyay, J. Appl. Phys. **77**, 5924 (1995).
[8] N. Telang, and S. Bandyopadhyay, Phys. Rev. **B48**, 18002 (1993); N. Telang, and S. Bandyopadhyay, Appl. Phys. Lett. **62**, 3161 (1993).
[9] H. Ishihara, and K. Cho, Phys. Rev. **B42**, 1724 (1990).

# Giant dipole effect and second-harmonic generation in quantum wires biased with a magnetic field

A. Svizhenko
*Department of Electrical Engineering, University of Notre Dame, Notre Dame, Indiana 46556*

A. Balandin[a] and S. Bandyopadhyay[b]
*Department of Electrical Engineering, University of Nebraska, Lincoln, Nebraska 68588*

(Received 7 November 1996; accepted for publication 17 March 1997)

We have theoretically studied giant dipoles associated with transitions between magneto-electric subbands in a quantum wire subjected to a transverse magnetic field. The strengths of these dipoles and their resonant frequencies can be varied with the magnetic field which then allows one to tune the emission wavelength of these transitions. The large magnitude of the dipole moments also leads to a strong second-harmonic component of the dielectric susceptibility that can be utilized for nonlinear optical applications such as second-harmonic generation, limiting, mixing, optical switching, etc. © *1997 American Institute of Physics.* [S0021-8979(97)02612-1]

## I. INTRODUCTION

Direct intraband transitions between the quantized states (subbands) of the conduction band in a quantum well is a well-researched topic.[1] It has been shown both experimentally and theoretically that such transitions have very large dipole moments and narrow bandwidths. Strong infrared absorption, associated with transitions between the lowest two electronic subbands in a GaAs quantum well, was observed long ago by a number of experimental groups.[2] Recently, population inversion between the second and third subbands of a quantum well has been established unambiguously and has led to demonstration of the celebrated quantum cascade laser.[3] The energy separation between the subbands in a quantum well or wire can be varied by an external magnetic field which then allows one to realize a continuously tunable laser or light-emitting-device. Moreover, the field can induce forbidden transitions that make additional frequency ranges accessible, thus permitting flexible device design.

Another potential use of magnetic field biasing of quantum wells or wires is in nonlinear optics. Nonlinear optical properties stem from higher order dielectric susceptibilties. Specifically, the second-order susceptibility $\chi^{(2)}$ is responsible for such phenomena as mixing and second-harmonic generation. It is well known that even-order susceptibilities vanish in structures with inversion symmetry. Consequently, finite second-order susceptibilities can be obtained in such structures only if the inversion symmetry of the conduction-band potential is broken either by an external electric field or by the intentional growth of an asymmetric well. Obviously, the former is the preferred method since an electric field can be continuously varied which allows one to tune the degree of symmetry breaking and the magnitude of $\chi^{(2)}$. This method, however, has a practical shortcoming. An electric field tilts the potential barriers of the well thereby allowing carriers to escape by tunneling or thermionic emission. This is especially serious in GaAs/AlGaAs systems where the barrier height is relatively small. It has been pointed out that the electronic states in a quantum confined system biased by a transverse electric field are never true bound states since the particles can always lower their energy by escaping from the well.[4] Therefore, these states have a finite lifetime, which broadens the transitions.

To overcome this shortcoming, one can adopt magnetostatic biasing. A magnetic field can break inversion symmetry without tilting potential barriers and promoting carrier escape. A transverse magnetic field applied to a quantum wire exerts a Lorentz force on an electron moving along the wire. As a result, its wave function (in any subband) will be skewed towards one edge of the wire. This skewing does not tilt potential barriers to first order (the barriers may tilt slightly because of a second-order effect associated with space charges and the self-consistent (Hall) electric field). However, it effectively breaks inversion symmetry since it causes a net charge to accumulate at either edge of the wire (the charges at the two edges have opposite signs as in the classical Hall effect). This leads to a nonvanishing even-order susceptibility in a symmetric structure. The skewing has another subtle effect. The degree to which the wave function is skewed is *different in different subbands* since an electron has different kinetic energies and hence experiences different Lorentz forces in different subbands. As a result, transitions between subbands whose wave functions have the same parity — which are forbidden without a magnetic field — are now allowed since the parities are altered by different amounts in different subbands by the different degrees of skewing.[5]

This article is organized as follows. In Section II, we describe the theoretical formulation, followed by results. Finally, in Section IV, we present the conclusions.

## II. THEORY

We consider a quantum wire as shown in Fig. 1 with a magnetic field applied along the $z$ direction. The thickness

[a]On leave from the University of Notre Dame.
[b]Electronic mail: bandy@engrssl.unl.edu

 0021-8979/97/81(12)/7927/7/$10.00 © 1997 American Institute of Physics 

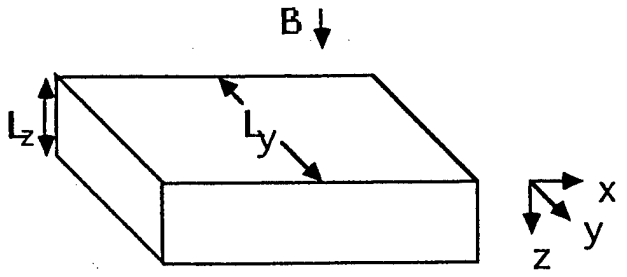



FIG. 1. An electron waveguide (quantum wire) subjected to a magnetic field along the $z$ axis. The width of the wire is much larger than the thickness.

along the $z$ direction is so small (and consequently the subband separation in energy is so large) that, for the range of photon energies considered, an electron cannot be excited (by real transition) into a subband which has more than two nodes along the $z$ direction. Such a transition will not be accessible in energy. This restriction, coupled with the fact that a magnetic field does not affect the $z$ component of the electron wave function, allows us to drop the $z$ component from further consideration. The width of the wire along the $y$ direction is however large enough that subbands with more than two nodes along the $y$ direction are accessible in energy.

In the framework of the envelope function approximation (EFA), an electron wave function can be written as the product of a Bloch wave function, periodic with the atomic lattice spacing, and an envelope wave function, describing the nonperiodic behavior. Consequently, the wave function of an electron for a given wave vector $k$ along the $x$ direction, in the $n$th magnetoelectric subband, at a magnetic field $B$ can be written as

$$\Phi_n(x,y,k,B,t)=\Psi_n(x,y,k,B)u_n(x,y,z,k)e^{-jE_n(k,B)t/\hbar}, \tag{1}$$

where $\Psi_n(x,y,k,B)$ is an envelope function, $u_n(x,y,z,k)$ is a Bloch function of a conduction band and $E_n(k,B)$ is the dispersion relation of the $n$th magnetoelectric subband at a flux density $B$. The Bloch wave functions are assumed to be $s$ states which is the usual case for semiconductors where $\hat{J}=1/2$ for the conduction band.

The envelope function can be further decomposed into a plane wave along the unconfined $x$ direction and a confined component along the $y$ direction. Thus,

$$\Psi_n(x,y,k,B)=\chi_n(y,k,B)e^{jkx}. \tag{2}$$

Using the electric dipole approximation, we can write the matrix element of photoinduced intersubband transitions within the conduction band as[6]

$$d_{f,i}(k,B)=e\int \chi_f(y,k,B)\hat{\eta}\cdot\vec{r}\chi_i(y,k,B)d\vec{r} \times \int u_f^*(x,y,z,k)u_i(x,y,z,k)d\Omega, \tag{3}$$

where $d\Omega$ is a volume element, $\hat{\eta}$ is the unit vector along the direction of the incident photon polarization, $\vec{r}=x\vec{a}_x+y\vec{a}_y$ is the two-dimensional radius vector, and subscripts $i,f$ stand for initial and final states respectively. The exponential term of Eq. (2) is not present in Eq. (3) since, for photoinduced transitions ($k_f=k_i$), the product of the exponential function and its complex conjugate are exactly unity. The volume overlap of the Bloch functions is also unity for $s$ states with the same wave vector. Now, if we assume that the incident light is polarized along the $y$ direction so that $\hat{\eta}=\hat{a}_y$, the above equation simplifies to

$$d_{f,i}(k,B)=e\langle\chi_f|y|\chi_i\rangle=e\int_{-W/2}^{W/2} y\chi_f(y,k,B)\chi_i(y,k,B)dy, \tag{4}$$

where $W$ is the width of the quantum wire along the $y$ direction.

One should note here that if there is no magnetic (or electric) field applied, the envelope functions $\chi_i$ are just particle-in-box states and the dipole moment in Eq. (4) is nonzero only for the transitions between subband states of opposite parity. For a symmetric square potential well, these dipole elements (between any two states $n$ and $m$) are independent of the wave vector $k$ and can be found analytically[1] by evaluating the integral in Eq. (4).

$$d_{f,i}=e\langle\chi_f|y|\chi_i\rangle=eW\frac{8}{\pi^2}\frac{mn}{(m^2-n^2)^2},$$

$$\text{if } n \text{ and } m \text{ have opposite parity}$$

$$=0, \text{ otherwise.} \tag{5}$$

However, when a magnetic field is applied, the skewing of the wave functions changes the integral in Eq. (4) and alters the selection rules. Generally, the skewing causes three effects. First, it makes the dipole moment depend on the wave vector $k$ (since the degree of skewing depends on $k$). Second, it reduces the dipole moment for transitions between states of opposite parity [since the integral in Eq. (4) decreases], and third, it allows forbidden transitions between states of the same parity [since the integral in Eq. (4) no longer vanishes for states of the same parity].

It is clear from Eq. (4) that, to calculate the dipole moments in the presence of a magnetic field, all we need to compute are the wave functions $\chi_{f,i}(y,k,B)$ at a given magnetic field $B$, for given magnetoelectric subbands $f$ and $i$, and for a given wave vector $k$. This is achieved via a numerical (finite difference) solution of the Schrödinger equation following the prescription of Ref. 7. Once this is done, we can calculate the dipole moment in Eq. (4) for any chosen intersubband transition at any chosen magnetic field and for any chosen wave vector.

In the limit of high magnetic fields, when the magnetic length $l(=\sqrt{\hbar/eB})\ll W$, one can again obtain an analytical expression for the dipole moment $d_{f,i}$. In this case, the magnetostatic confinement predominates over electrostatic confinement and the envelope functions $\chi_n(y,k,B)$ can be approximated by harmonic-oscillator wave functions:

$$\chi_n(y,k,B)\equiv\chi(y-y_k,B)$$

$$=N_nH_n(\alpha,y-y_k)e^{-\frac{1}{2}\alpha^2(y-y_k)^2}, \tag{6}$$

where $N_n=(\alpha/\pi^{1/2}2^2n!)^{1/2}$ is a normalization constant, $H_n(\alpha,y)$ is the $n$th Hermite polynomial, $y_k=\hbar k/eB$, and

$$\alpha=\sqrt{\frac{eB}{\hbar}}=\frac{1}{l}. \tag{7}$$

In order to evaluate the integral in Eq. (4) analytically, we extend the limits of integration to infinity assuming that the wave function tail is negligible at the boundaries of the wire (i.e. at $y=\pm W/2$). This is a very reasonable assumption in a high confining magnetic field. The resulting analytical expression for the dipole moment is

$$d_{f,i}(B)=e\langle\chi_f|y|\chi_i\rangle=el\left(\frac{n+1}{2}\right)^{1/2}, \quad \text{if } m=n+1 \tag{8}$$

$$=el\left(\frac{n}{2}\right)^{1/2}, \quad \text{if } m=n-1$$

$$=0, \text{ otherwise.}$$

The physical significance of the two analytical limits, $B\to 0$ and $B\to\infty$, is obvious. At zero field, the dipole is determined by the width of the wire $d_{f,i}\sim eW$, and at the high field limit it is determined by the magnetic length $d_{f,i}\sim el$. This is what one would expect intuitively. At zero field, the dipole is confined electrostatically with the wire width being a measure of this confinement while at high magnetic field, the dipole is confined magnetostatically and the magnetic length is the corresponding measure of this confinement.

## III. RESULTS

### A. Intraband dipoles

We now present results of our calculations. The physical parameters used for the numerical calculations correspond to a GaAs quantum wire with relative dielectric constant $\epsilon_r=12.9$, and effective mass $m_e=0.067m_0$ where $m_0$ is the free electron mass.

In Fig. 2, we show the dependence of the dipole moment $d_{f,i}(k,B)$ for three transitions ($e1$-$e2$, $e2$-$e3$, and $e1$-$e3$) on the wave vector $k$ when a magnetic field of 1 T is applied (following usual practice, the transitions are numbered by the subband indices). The dipoles corresponding to transitions between states of opposite parity ($e1$-$e2$ and $e2$-$e3$) have maxima at $k=0$ and then decrease with increasing wave vector. This can be easily understood as follows. At zero wave vector (no translational velocity) these states do not experience any Lorentz force and hence the wave functions are not skewed. As the wave vector $k$ increases, the translational velocity and the Lorentz force experienced increase. Consequently, the envelope wave functions are skewed more and more and the dipole moment decreases. Real transitions between states of the same parity are forbidden at zero magnetic field, but at a finite magnetic field, they are forbidden only at $k=0$ when there is no translational velocity and no Lorentz force to skew the wave functions. With increasing $k$, the wave functions are increasingly skewed and the dipole moment of forbidden transitions increases. In our chosen

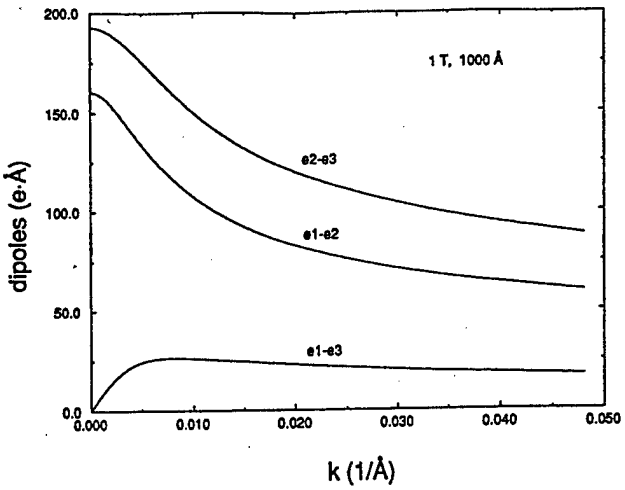



FIG. 2. Dipole moments for various intersubband transitions as functions of the propagating wave vector $k$ for a magnetic flux density of 1 T. At zero translational velocity ($k=0$) the dipole of transition $e1$-$e3$ vanishes. The GaAs quantum wire is 1000 Å wide.

prototype wire, $d_{e1-e3}$ reaches a maximum of 28 $e$-Å at $k=0.0051$ Å and then decreases gradually ultimately reaching zero. This intriguing *nonmonotonic* dependence on $k$ is explained later on. However, at this point, it is interesting to note that a fairly large forbidden dipole moment of ~30 $e$-Å can be achieved in realistic structures at a moderate magnetic field of 1 T.

Fig. 3 presents the dipole moments for the same transitions as a function of magnetic flux density. The propagation wave vector $k$ is chosen to be 0.01/Å. At zero magnetic field, a nonvanishing dipole matrix element occurs only for transitions between states of opposite parity ($e1$-$e2$, $e2$-$e3$) as expected from Eq. (5). This equation also allows us to estimate the strengths of these zero-field dipoles to be 180 $e$-Å for $e1$-$e2$ and 195 $e$-Å for $e2$-$e3$ transitions. As we can see from Fig. 3, these values are in excellent agreement with our

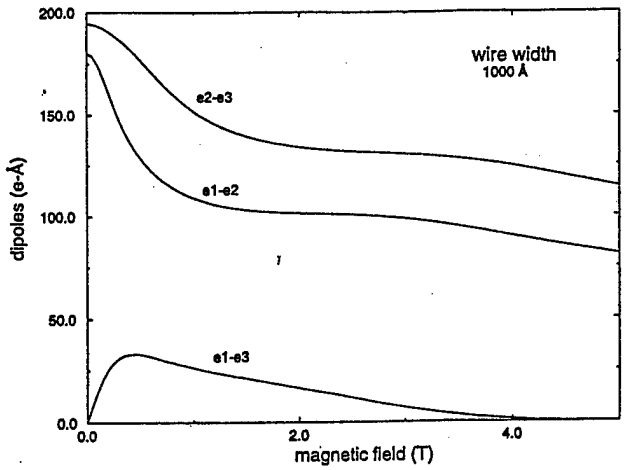



FIG. 3. The dipoles of three intersubband transitions as functions of the applied magnetic field. The dipole $d_{e1-e3}$ peaks at a magnetic flux density of 0.3 T. The wire width is the same as that in Fig. 2.

numerical results. From the analytical expression in Eq. (8), we can estimate the strength of the $e1$-$e2$ dipole to be 66 $e$-Å at a magnetic flux density of 15 T. This number also agrees with our numerical result. The $e1$-$e3$ dipole vanishes at both zero field (because of the spatial symmetry of the particle-in-a-box states) and at high fields because of the symmetry of the Landau states or Hermite polynomials. This behavior is consistent with Eqs. (5) and (8). Only at intermediate fields, when the wave functions of the subbands are a hybrid between particle-in-a-box states and Hermite polynomials (and thus "nonsymmetric" in space), is this transition allowed. This immediately tells us that $d_{e3-e1}$ must have a *nonmonotonic* dependence on the magnetic flux density $B$ and indeed it does.

Let us now examine the nonmonotonic behavior of $d_{e3-e1}$ more closely. This transition is forbidden at zero field since the wave functions of the first and third subbands have the same parity. At low magnetic fields, the parities are altered by the skewing of the wave functions and consequently $d_{e3-e1}$ is no longer zero but increases with the magnetic field. It reaches a maximum of about 30 $e$-Å and then decreases. This latter decrease is related to the following effect. For a fixed wave vector $k$, a sufficient increase in the flux density $B$ forces the traversing states ("skipping orbits" or "edge states") to condense into closed cyclotron orbits (Landau levels) that are no longer skewed by the magnetic field to the wire edge since they have no translational velocity and hence no Lorentz force. While edge states have a skewed wave function that is not symmetric in space, cyclotron orbits have a wave function that is symmetric about the orbit center $y_k$. Note that $y_k$ depends only on $k$ and $B$. Therefore, at a fixed $k$, the wave functions of the first and third Landau levels are symmetric about a *common* center. Whenever this kind of symmetry holds, $d_{e3-e1}$ vanishes. Therefore, the dipole moment $d_{e3-e1}$ decreases gradually to zero at high magnetic field with the onset of Landau condensation.

The same physics can be elucidated in a different way by considering the energy versus wave vector relation in Figs. 4(a) and 4(b) which show the dispersion of the first and third magneto-electric subbands respectively.

At $B=0$, velocity (slope of the curves) at $k=0.01$/Å are nonzero for both the $e1$ and $e3$ subbands. However, the Lorentz force is zero because $B=0$ and hence $d_{e1-e3}=0$. At $B=5$ T, the group velocities for the two subbands are still nonzero and the Lorentz force is finite resulting in skewing of wave functions and a nonvanishing value of $d_{e1-e3}$. At $B=10$ T, the group velocities at $k=0.01$/Å are zero in both subbands indicating that the corresponding states have undergone Landau condensation. In this case, the Lorentz force (for skewing) is again zero and the dipole moment $d_{e1-e3}$ vanishes once more. The crucial point to note is that the Lorentz force $e\vec{v}\times B$ can vanish in two different ways: (i) $B=0$, and (ii) $\vec{v}=0$. These two conditions are met at zero and very high magnetic fields. As a result, the dipole moment $d_{e1-e3}$ exhibits a nonmonotonic behavior in magnetic field. One can ask why the same physics does not cause nonmonotonicity in the $e1$-$e2$ and $e2$-$e3$ curves. It is not clear *a priori* that nonmonotonicity cannot occur (indeed

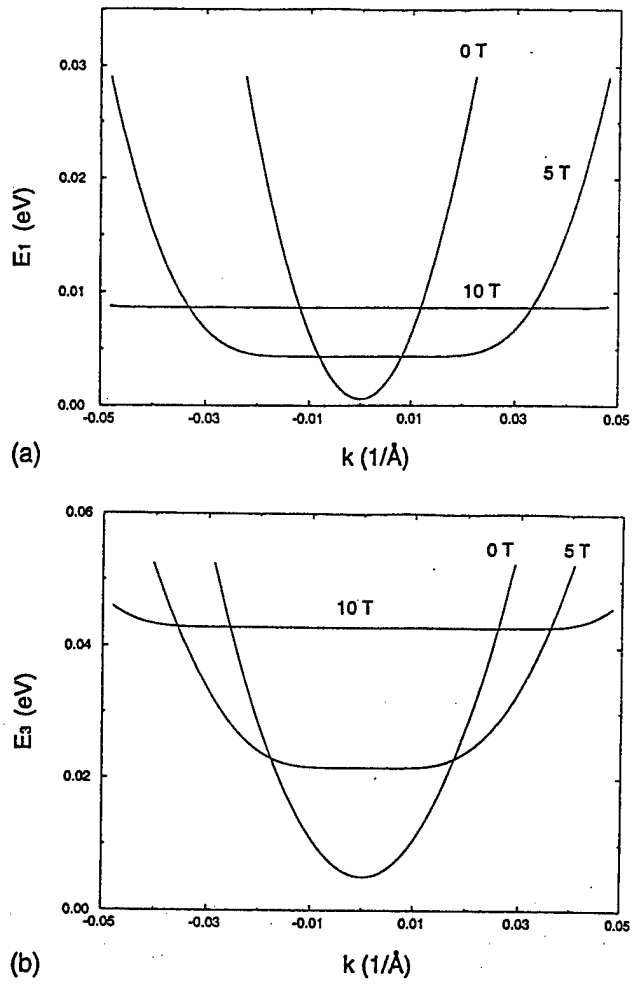



FIG. 4. Energy vs wave vector relation of electrons in (a) the first subband and (b) in the third subband of a 1000 Å wide quantum wire. The wave vector is along the free propagation direction. The results are shown for three values of a magnetic field. The energy is calculated from the bulk conduction band edge and the confinement energy for the z-direction is assumed to be zero.

there are regions of inflexion in the two curves). However, the point to note is that Landau condensation causes recovery of the wave function symmetry (or antisymmetry), *but does not restore the original zero-field wave functions*. This is shown in Fig. 5 where we show the wave functions in the $e1$ subband at 0 and 10 T. Both wave functions are "symmetric" in space, but they are otherwise vastly different since the magnetostatic confinement squeezes the wave functions binding them in cyclotron orbits.

The nonmonotonicity in the wave vector dependence of $d_{e1-e3}$ in Fig. 2 has a similar origin. As $k$ is increased, the *relative* skewing between the wave functions in $e1$ and $e3$ subbands change nonmonotonically causing the nonmonotonicity seen.

The process described above is illustrated in Figs. 6(a)–6(c), where we present wave functions of two electronic states ($e1$ and $e3$) for three values of magnetic flux density. At zero magnetic field the wave functions are symmetric about the center of the wire and dipole transition $d_{e3-e1}$ is forbidden [Fig. 6(a)]. At low magnetic field the wave functions are skewed to the edge of the wire ["edge states" Fig.

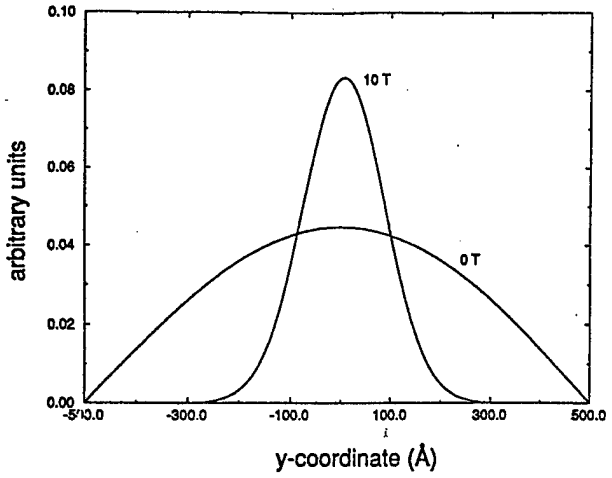



FIG. 5. The $y$ component of the electron envelope function for the first subband at a magnetic flux densities of 0 and 10 T.

6(b)] and the spatial symmetry is broken for both states. Dipole transition $d_{e3-e1}$ is now allowed. It is important to note here that the symmetry breaking skewing of the wave functions is caused by the simultaneous presence of a magnetic field and the electrostatic potential barriers at the edges of the quantum wire. At higher magnetic fields, when the magnetic length is smaller than the wire width, the electrons do not "feel" the potential barriers at the edges of the wire as they undergo complete Landau condensation and execute cyclotron motion with a radius much smaller than the width of the wire. In this case, the wave function symmetry is essentially restored [Fig. 6(c)] although the wave functions are now symmetric about a point that is not at the center of the wire. Nonetheless, what is important is that both wave functions are symmetric about the same point. Consequently, the $d_{e3-e1}$ transition vanishes. The simultaneous presence of both electrostatic confinement and magnetostatic confinement is therefore necessary for wave function skewing, formation of edge states and the observation of forbidden transitions.

## B. Second-harmonic generation

It is well known that in systems with inversion symmetry there can be no second order nonlinearity.[8] However, in systems without inversion symmetry, the lowest order optical nonlinearity is of the second order and is expressed by

$$\vec{P}^{(2)}(\vec{k},\omega)=\chi^{(2)}(\omega;\omega_1,\omega_2)\vec{E}_1(\vec{k_1},\omega_1)\vec{E}_2(\vec{k_2},\omega_2), \quad (9)$$

where $\vec{P}$ is the polarization caused by two electric fields $\vec{E}_1$ and $\vec{E}_2$ that are associated with the electromagnetic fields of either two frequency components of the same light beam or two different coherent beams with frequencies $\omega_i$ and wave vectors $\vec{k}_i$. The frequencies and wave vectors obey the energy and momentum conservation laws

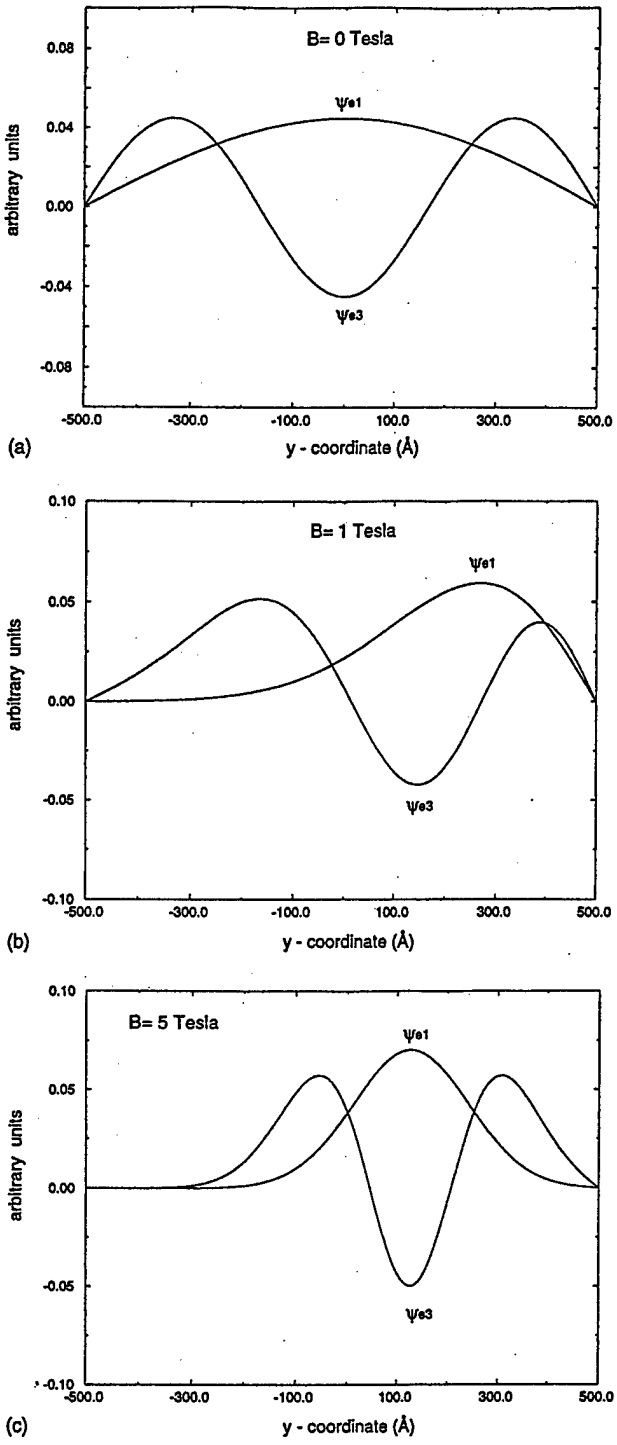



FIG. 6. The $y$ component of the electron envelope functions for the first and third electronic subbands. The results are shown for cases when (a) no magnetic field is present, (b) when a weak magnetic field is present and, finally, (c) when a strong magnetic field is present.

$$\hbar\omega=\sum_i \pm\hbar\omega_i, \qquad \hbar\vec{k}=\sum_i \pm\hbar\vec{k}_i. \quad (10)$$

It is obvious that the third-ranked tensor $\chi^{(2)}$ will vanish in any structure with inversion symmtery. A quantum con-

fined structure may lack inversion symmetry for two main reasons. (i) The semiconductor material by its intrinsic chemical and crystalline structure may lack inversion symmetry,[9] and this is the case in most III–V, II–VI, and I–VII compounds along certain crystallographic directions. (ii) The quantum confining potential well may be asymmetric (e.g. triangular potential well, asymmetric double square well potential, etc.). In the first case, the asymmetry is related to the intracell charge asymmetry and is not affected by the confinement since the latter extends over several unit cells. In the second case, the asymmetry is artificially imposed and therefore can be engineered. It clearly depends on the confining potential and hence an applied electric field can alter the potential and change the degree of symmetry breaking.

In the present work we restrict ourselves to the second case and do not consider intrinsic second-order nonlinearities of GaAs which are actually quite large (the nonlinear susceptibility of bulk GaAs is $\chi^{(2)}_{14}=3.8\times10^{-10}$ m/V[10]). As mentioned before, we avoid an electric field since it promotes carrier escape and we consider a magnetic field instead. Although a magnetic field does not affect the potential to first order, the simultaneous action of *symmetric* electrostatic potential and an external magnetic field may lead to the *uneven charge distribution* along the width ($y$ axis) of the wire caused by different degrees of skewing of the wave functions. Because of this reason, it is possible to break the inversion symmetry in a symmetric quantum well or wire with a magnetic field alone. This approach is superior to applying a transverse electric field since the latter will tilt the confining potential wells thereby promoting carrier escape from the well by either tunneling or thermionic emission.

The large magnitude of the dipole moments associated with otherwise forbidden transitions between subbands of the same parity and their sensitivity to the biasing magnetic field opens up the possibility of second-harmonic generation (SHG) that can be controlled by the magnetic field. In order to evaluate the magnitude and dependences of SHG on the biasing field and wire geometry, we calculate the second-order susceptibility using the formula[11]

$$\chi^{(2)}_{\mu\alpha\beta}(-\omega_\sigma;\omega_1;\omega_2)=\frac{Ne^3}{\epsilon_0 2\hbar^2}\hat{S}_T\sum_{abc}\rho_o(a)\times\left[\frac{d^\mu_{ab}d^\alpha_{bc}d^\beta_{ca}}{(\Omega_{ba}-\omega_1-\omega_2)(\Omega_{ca}-\omega_2)}\right], \quad (11)$$

where $N$ is the concentration (number density) of conduction electrons, $\hbar\Omega_{\alpha\beta}\equiv\hbar\Omega_{\alpha\beta}(B,W)$ is the energy spacing between $\alpha,\beta$ subbands that depends on the applied magnetic field and wire width, $d_{mn}\equiv d_{mn}(B,W)$ is a dipole element calculated using Eq. (4), and $\omega_\sigma$ is defined to be $\omega_\sigma=\omega_1+\omega_2$. The total symmetrization operation $\hat{S}_T$ indicates that the expression that follows it is to be summed over all six permutations of the pairs $(\mu,-\omega_\sigma)$, $(\alpha,\omega_1)$, $(\beta,\omega_2)$. Since $\hat{S}_T$ involves a summation over all possible permutations, it is clear that $\chi^{(2)}_{\mu\alpha\beta}(-\omega_\sigma;\omega_1;\omega_2)$ is invariant under any of them. For simplicity, the Fermi distribution $\rho_o(a)$ was assumed to be unity.

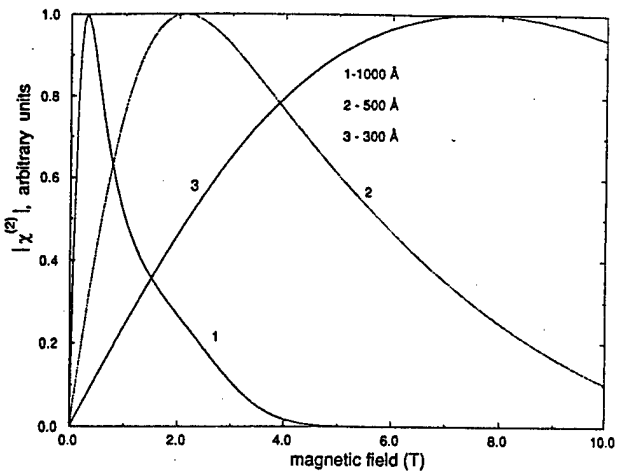



FIG. 7. Second-order susceptibility as a function of the biasing magnetic field. The peak values of the susceptibility are $13.2\times10^{-7}$ m/V, $1.5\times10^{-7}$ m/V and $3\times10^{-8}$ m/V for wire widths of 1000 Å, 500 Å, and 300 Å, respectively. The results are shown for the wave vector $k=0.01$/Å (fixed excitation frequency).

Eq. (11) is an approximation that applies only under the condition that all of the optical frequencies involved (operational frequencies $\omega_\sigma,\omega_1,\omega_2$) are removed far enough from the subband transition frequencies. It means that the medium is assumed to be transparent and loss free at all the relevant optical frequencies. This assumption can be relaxed by the introduction of transition damping factors into the expression in Eq. (11). In our study we are mainly interested in the effects of an applied magnetic field on the second-order susceptibility. Since these effects manifest themselves in Eq. (11) primarily via the magnetic field dependence of the dipole elements $d_{mn}\equiv d_{mn}(B)$, we did not include any damping constants and associated finite linewidths of the electronic states. One should also note here that Eq. (11) is strictly correct only for dilute media. In this case, one can write $\chi^{(2)}=N\alpha^{(2)}$ with $\alpha^{(2)}$ being the second-order nonlinear polarization. The above expression is valid only under moderate excitation.

In Fig. 7, we present normalized values of $\chi^{(2)}$ as a function of magnetic field for three different wire widths and a fixed value of the wave vector $k$ (fixed excitation frequency). The operational frequencies $\omega_1=\omega_2$ are chosen for a $CO_2$ laser. For wide ranges of magnetic flux densities ($B<20$ T) and wire widths (100 Å $<W<$1000 Å), these frequencies are removed far enough from the subband transition frequencies $\Omega_{\alpha\beta}(B,W)$. As long as the latter is true, the $\chi^{(2)}$ dependence on magnetic field is governed mainly by dipole elements $d_{mn}$. Consequently, the $\chi^{(2)}$ curve for $W=1000$ Å peaks at the same value of a magnetic flux density ($B=0.3$ T) as the $e1$-$e3$ dipole curve of Fig. 3. The magnetic flux density at which $\chi^{(2)}$ reaches its maximum increases with decreasing wire width. This happens because it takes a higher magnetic field to condense electronic states into cyclotron orbits (Landau condensation) when the electrostatic confinement is stronger (narrower wires).

Fig. 8 shows the dependence of the normalized values of $\chi^{(2)}$ on wire width for three different values of a magnetic

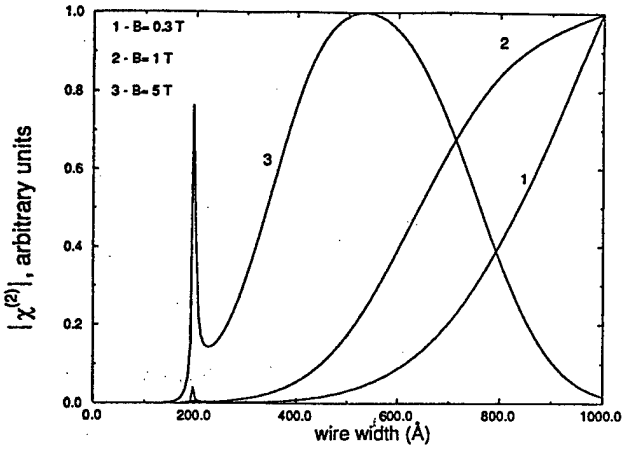



FIG. 8. Second-order susceptibility as a function of the wire width for three values of the biasing magnetic field. The maximum values of the $\chi^{(2)}$ curves are the same as in Fig. 7. The narrow peaks at a wire width of ~200 Å are due to resonances occurring when either $\Omega_{ba}=\omega_1+\omega_2$ or $\Omega_{ca}=\omega_2$.

field and a fixed value of the wave vector. For weak magnetic field of 0.3 T, the $\chi^{(2)}$ curve increases monotonically with increasing wire width. This happens because $d_{m,n}$ is proportional to the wire width $W$ (see Eq. (5) which is valid at zero field). The magnetic field is obviously not strong enough for the onset of Landau condensation. At a moderate magnetic flux density of 1 T, we can observe some saturation features, and for a strong magnetic field of 5 T, the curve is nonmonotonic, rolling down to almost zero for the wire width of 1000 Å. The physics underlying the difference in the behaviors of the three curves is essentially the same as that responsible for the features in Fig. 7. At small values of wire width ($W\approx 170$ Å), there is an additional peak in the $\chi^{(2)}$ curve. This peak is a manifestation of the fact that $\Omega_{ba}$ has become comparable to the operational frequencies, i.e. $\Omega_{ba}\approx\omega_1+\omega_2$.

In our numerical calculations we have used $N=10^{17}$ $cm^{-3}$. For this dilute concentration, high density effects such as screening and bandgap renormalization are not important and Eq. (11) is strictly valid. In fact, Ref. 1 demonstrated excellent agreement between theory and experiment without accounting for any high density effect even though the carrier concentration in that study was $N=5\times10^{17}/cm^3$. Therefore, we believe that high density effects are not significant in this regime.

The peak value of the second-order susceptibility for a wire width of 500 Å is $\chi^{(2)}=1.5\times10^{-7}$ m/V (the absolute magnitudes of the peak values for various wire widths are given in the caption of Fig. 7). For comparison, the nonlinear susceptibility of electric field biased GaAs quantum wells ($W=92$ Å), calculated theoretically and measured experimentally in Ref. 1, was $\chi^{(2)}=2.4\times10^{-8}$ m/V for an electric field of 36 kV/cm. This shows that relatively weak magnetic fields in quantum wires can produce similar magnitudes of $\chi^{(2)}$ as rather strong electric fields in quantum wells. Unfortunately, to our knowledge, there is no theoretical or experimental result available for either electric field biased quantum wires or magnetic field biased quantum wells so a direct comparison is not possible. Nonetheless, it is obvious that magnetic field biased quantum wires provide a very attractive alternative to other means of producing large $\chi^{(2)}$ values. In fact, the largest value of $\chi^{(2)}$ (obtained at a magnetic flux density of 2 T) in a magnetic-field-biased quantum wire is found to be three orders of magnitude higher than what can be achieved in bulk GaAs.

## IV. CONCLUSION

We have theoretically studied the giant dipole effect in magnetic-field-biased semiconductor quantum wires. The dipoles are associated with transitions between magneto-electric subbands within the conduction band, some of which are forbidden in the absence of the magnetic field. The resonant frequencies of these transitions can be tuned by the magnetic field which allows the realization of externally tunable intersubband lasers. We have also studied the possibility of second-harmonic generation in a quantum wire biased with a magnetic field and find a strong second-harmonic component of the susceptibility. This may have important applications in nonlinear optics.

## ACKNOWLEDGMENT

This work is supported by the U.S. Army Research Office under Contract Nos. DAAH04-95-1-0586 and DAAH04-95-1-0527.

[1] M. M. Fejer, S. J. B. Yoo, R. L. Byer, A. Harwit, and J. S. Harris, Phys. Rev. Lett. **62**, 1041 (1989), and references therein; L. C. West and S. J. Eglash, Appl. Phys. Lett. **46**, 1156 (1985); for an overview of current work on intersubband transitions, see, for example, *Quantum Well Intersubband Transition Physics and Devices*, NATO ASI Series E270, edited by H. C. Liu, B. F. Levine, and S. Y. Anderson (Academic, Dordrecht, 1994).

[2] A. Sa'ar, I. Grave, N. Kuze, and A. Yariv, *Nonlinear Optics: Materials, Phenomena and Devices*, (IEEE, New York, 1990) 113; B. F. Levine, R. J. Malik, J. Walker, K. K. Choi, C. G. Bethea, D. A. Kleinman, and J. M. Vandenberg, Appl. Phys. Lett. **50**, 273 (1987).

[3] See, for example, Manfred Helm, Semicond. Sci. Technol. **10**, 557 (1995).

[4] E. J. Austin and M. Jaros, Phys. Rev. B **31**, 5569 (1985).

[5] A. Balandin and S. Bandyopadhyay, J. Appl. Phys. **77**, 5924 (1995); A. Balandin, Ph.D. dissertation, University of Notre Dame, 1996.

[6] C. Weisbuch and B. Vinter, *Quantum Semiconductor Structures: Fundamentals and Applications* (Academic, Boston, 1991).

[7] S. Chaudhuri and S. Bandyopadhyay, J. Appl. Phys. **71**, 3027 (1992).

[8] See, for example, J. M. Hvam in *Nonlinear Spectroscopy of Solids: Advances and Applications*, edited by B. Di Bartolo and B. Bowlby (Plenum, New York, 1994), pp. 91–149.

[9] This lack of symmetry is due to the transfer of a valency charge from one atom to the other bond-forming atom and resultant uneven charge distribution along the bond axis.

[10] *Handbook of Lasers*, edited by R. J. Pressley (Chemical Rubber, Cleveland, 1971), p. 504.

[11] See, for example, P.N. Butcher and D. Cotter, *The Elements of Nonlinear Optics* (Cambridge University Press, Cambridge, 1990); A. Yariv, *Quantum Electronics* (Wiley, New York, 1989); or Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley, New York, 1984).

# Collective Computational Activity in Self-Assembled Arrays of Quantum Dots: A Novel Neuromorphic Architecture for Nanoelectronics

Vwani P. Roychowdhury, *Member, IEEE,* David B. Janes, *Member, IEEE,*
Supriyo Bandyopadhyay, *Senior Member, IEEE,* and Xiaodong Wang

*Abstract*—**We describe a new class of nanoelectronic circuits which exploits the charging behavior in resistively/capacitively linked arrays of nanometer-sized metallic islands (quantum dots), self-assembled on a resonant tunneling diode, to perform neuromorphic computation. These circuits produce associative memory effects and realize the additive short-term memory (STM) or content addressable memory (CAM) models of neural networks without requiring either large-area/high-power operational amplifiers, or massive interconnectivity between devices. Both these requirements had seriously hindered the application of neural networks in the past. Additionally, the circuits can solve NP-complete optimization problems (such as the traveling salesman problem) using single electron charge dynamics, exhibit rudimentary image-processing capability, and operate at room temperature unlike most quantum devices. Two-dimensional (2D) processors, with a 100 × 100 pixel capacity, can be fabricated in an area of $10^{-8}$ cm$^2$ leading to unprecedented functional density. Possible routes to synthesizing these circuits, employing self-assembly, are also discussed.**

## I. Introduction

It is widely believed in the solid state device community that conventional strategies for integrating devices on a chip will be impractical for nanoelectronic devices because of their small size (~100 nm), low power handling capacity and low gain. Accordingly, many new proposals have been advanced, articulating primarily, visions of future architectures for nanoelectronic circuits [1]–[10]. Some of these proposals, which contemplate building Boolean logic circuits using Coulomb interaction between bistable charge polarizations in quantum dots are flawed [5]. They have overlooked the fact that bistability is a necessary, but not a sufficient condition for implementing Boolean logic circuits. Individual logic devices must also have *isolation* between the input and output (as in a conventional transistor) so that the input drives the output and not the reverse. Coulomb interaction between two identical charge polarizations is *reciprocal* so that it is impossible to distinguish the "input" polarization from the "output" polarization (i.e., there is no isolation between the input and output). In other words, the output influences the input just as much as the input influences the output! Consequently, logic signal cannot propagate *unidirectionally* from the input to the output, or from one stage to the next, leading to total failure [6], [7], [11]. The failure occurs because the input cannot uniquely (and predictably) determine the output. This problem is pathological in many proposed schemes of nanoelectronic architecture. Other proposals, which advocate the use of *cellular automata* architectures with local connectivity [1], also have shortcomings; they offer no mechanism for loading the initial program into the cellular array of quantum devices [11]. As a result, they are also of questionable efficacy. Recently, a scheme, based on the precise phase-locking of single electron tunneling oscillations in capacitively coupled nanojunctions to realize the parametron computing model of von-Neumann and Goto, has been proposed [10]; but it seems to lack fault-tolerance. Finally, some recent proposals for dissipationless computing [4], although quite intriguing, can hardly be considered practical since they have no error-correcting capability. Added to all this is the fact that most of these schemes make unreasonable demands on materials and fabrication technology that cannot be met in the near future.

In [6], [8], and [9], we proposed a radically different scheme for nanoelectronic architecture that suffers from none of the drawbacks just discussed. It has a number of unusual features. First, we realized that any proposal, which requires extreme fabrication tolerance whereby billions of nanoelectronic devices must be fabricated on a wafer with nominally identical behavior, is unrealistic. There will always be *stochasticity* in a large-scale system (e.g., a complex computer or signal-processing chip) and it may be wiser to exploit this stochasticity to realize useful functions rather than strive against it. This is precisely what we do. Second, we insist on *room temperature* operation since cryogenic operation is impractical. Finally, we restrict attention to niche applications where the collective (and sometimes stochastic) activity of a large number of nanostructure devices, working in unison, gives rise to useful computational activity.

The architecture that we implement is inspired by standard models of neural networks. It is massively parallel and inherently fault-tolerant. Moreover, it has significant fabrication tolerance (a 100% variation in the size of an individual device is quite tolerable) and a great deal of noise immunity. The

Manuscript received April 29, 1996. This work was supported in part by the U.S. Army Research Office under Grants DAAH04-95-1-0586 and DAAL03-G-0144, and by the U.S. National Science Foundation under Grant ECS 9308814.

V. P. Roychowdhury, D. B. Janes, and X. Wang are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA.

S. Bandyopadhyay is with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556 USA.

Publisher Item Identifier S 0018-9383(96)07213-9.

0018–9383/96$05.00 © 1996 IEEE

relative insensitivity to size variation accrues from the fact that the size determines the capacitance of an individual device. A 100% variation in the size of an individual device will result in a similar variation in the capacitance and this does not affect the performance of the circuit as a whole very much because of the "collective" nature of the computation. Here the collective activities of all devices acting cooperatively matter, rather than the characteristic of a single device. The switching speed of every device is $\sim$1 ps and the power delay product can approach the fundamental thermodynamic limit of $kT\ ln\ 2$ [12] which is less than $10^{-20}$ Joules. Needless to say, all of these are very desirable features.

This paper is organized as follows. In the next section, we describe a prototype nanoelectronic network fashioned from a 2D array of metallic islands self-assembled on the surface of a double barrier resonant tunneling diode structure. This very simple system produces associative memory effects, realizes the short-term additive memory (STM) or content addressable memory (CAM) model of neural networks, performs 2D image processing and solves NP-complete combinatorial optimization problems. Examples of all this are provided in Section III. In Section IV, we discuss possible techniques for fabricating the network. Finally, in Section V, we provide the conclusions.

## II. A Self Assembled Array of Metallic Islands on a Resonant Tunneling Diode: A Neuromorphic Circuit

In the past, many proposals for nanoelectronic circuits advanced ideas that could not be implemented in practice using technology that is either presently available or is likely to be available in the near future. In contrast, the basic array of circuit nodes in our proposed system can be fabricated with relative ease. The primary technology required to produce our circuits is a technology to create a 2D periodic array of nanometer-sized metallic islands (with nearest neighbor interconnections) on a substrate whose current-voltage characteristic has a *nonmonotonic* nonlinearity. Individual islands can have a diameter between 10 and 100 nm depending on the application. Such a structure is shown in Fig. 1(a). The nearest-neighbor interconnects are implemented by organic molecular wires or some other suitable means. This is an established technology and the details of the fabrication will be given in Section IV. The substrate can be a resonant tunneling diode (RTD) whose conduction characteristic has the required nonmonotonic nonlinearity [13]. All our circuits utilize this basic structure. We assume that the molecular wires or other interconnections between the metallic islands can produce a variety of different network mechanisms (e.g., phase-coherent, or ohmic, or nonlinear, or single-electron, or rectifying transport, etc.) which lead to a rich plethora of circuit functions. Some of these possible network mechanisms are depicted in Fig. 1(b)–(d).

Electrically coupled arrays of nanometer scale metallic islands, comparable to that depicted in Fig. 1(a), have been fabricated in the past using self-assembly [14]. There is experimental evidence that one can establish inter-island conductive and capacitive links between the nanometer sized metallic islands with organic molecular wires [14]. Single electron charging effects and nonlinear transport have been demonstrated in these systems at room temperature [15]. Furthermore, self-assembled buckytubes may also realize the interconnects [16], [17]. Additionally, there is experimental evidence that phase-coherent links may prevail between metallic spheres which have been embedded in a porous dielectric matrix [18]. Therefore, a wide variety of transport mechanisms can be realized at present which provides an adequate degree of design flexibility.

The final requirement for realizing the complete system shown in Fig. 1(a) is a suitable scheme for reading and writing data. In the topologies described in Section III, each circuit node must be addressed individually. Present interconnect technology is not capable of providing individual connections at this density, although advances in nanometer scale patterning or self-assembly techniques may provide such capabilities in the near future. In the shorter term, related computational topologies with somewhat reduced functionality can be realized with all input/output ports and program nodes at the periphery of the chip. In that case, all input/output are provided to or retrieved from a few selected nodes at the periphery or edge of the chip. An architecture based on such "edge-driven" paradigm is represented by the general structure proposed in Fig. 1(a). A scheme that can implement logic functions using such a structure has been recently proposed. The interconnect problem is greatly simplified in the "edge-driven" case, since it is necessary to externally access only those islands that communicate with the external world. Of course, conventional bonding technology will not be adequate to access these islands because it does not possess the resolution needed to address individual quantum dots. In fact, typical bonding pads will probably consume the area occupied by about 10 000 islands. Therefore, we need a contact technology capable of virtually atomic resolution. The obvious choice is a scanning tunneling microscope (STM) tip which has been shown to be capable of such resolution [19]. They can also read the voltages on the island. STM tips were proposed for input/output data lines in [6] which visualized an array of tips micromachined into a wafer and permanently attached to input/output ports for data reading and writing. Attaching an STM tip to an individual dot may be accomplished using fine line direct-write lithography followed by metallization. This is a difficult step and significant developmental work should precede implementation. Recently, the operation of an integrated 100 micron size STM fabricated on a chip by standard lithography was demonstrated by Hitachi Research Laboratories [20] which is an important advancement in this direction. It should be noted that the size of the dots in our proposed structures is typically 10 nm so that once the STM tips are correctly attached, thermal drift or vibration is no longer a problem. Additionally, one should note that direct write lithography is being used only for making contacts to the chip and not for delineating the quantum dots or the interconnects between them. Therefore, the throughput will not be unacceptably slow even though direct write lithography is used. Recent work on arrays of scanning tunneling microscope tips, in which each tip has been individually addressed [21] lends further credence to this idea.

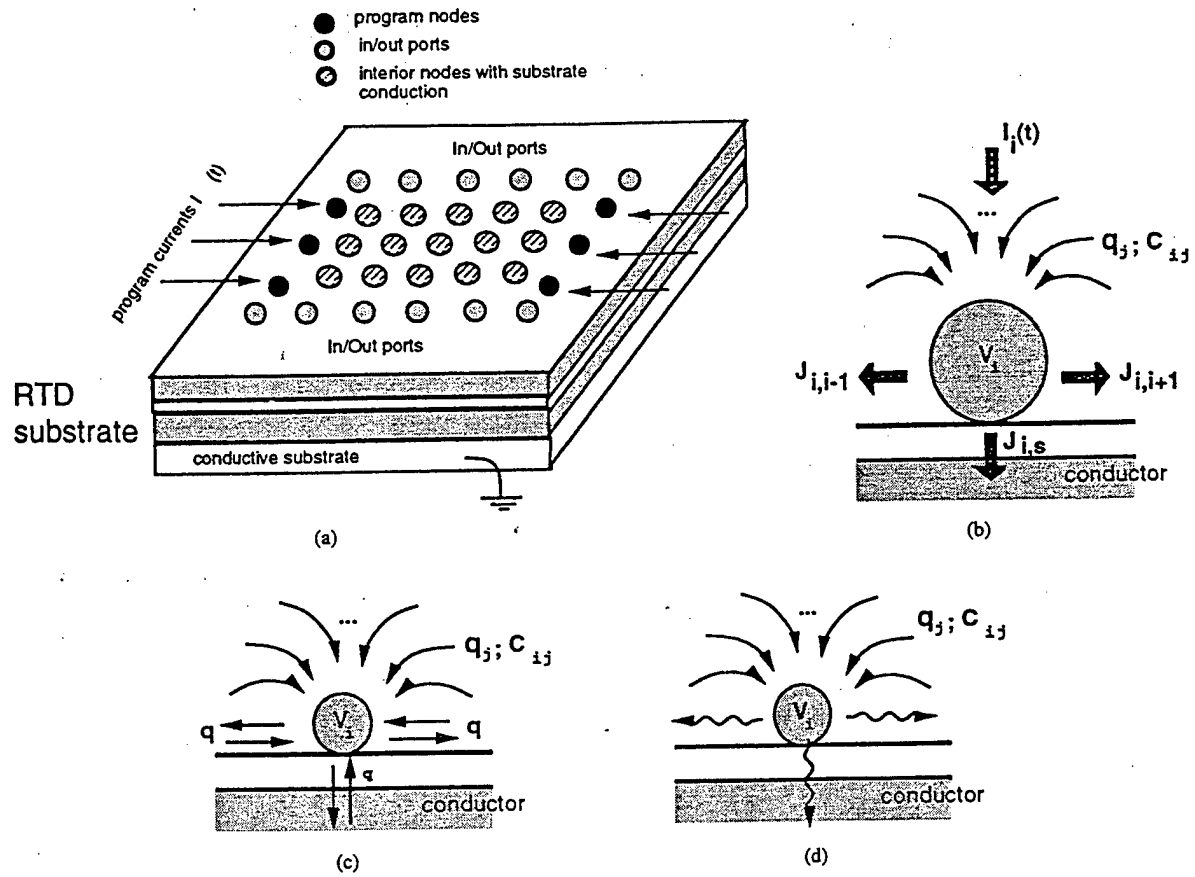



Fig. 1. A generic array of nanometer-sized metallic islands deposited on a resonant tunneling diode (RTD). It is assumed that all islands have direct conductive/capacitive links to their nearest neighbors established, for example, via organic molecular wires. These wires are not shown for the sake of clarity. A subset of the islands serve as program nodes that are driven by external current sources. Another subset of islands serve as input/output ports, and the remaining islands introduce complexity to the system through nonmonotonic nonlinear conductive link to the substrate; (b)–(d) Cross-sectional view showing island $i$, whose potential $v_i$ is influenced by the charges $q_j$ in the rest of the network through a capacitance matrix. Continuous charge dynamics (where electric charge is a classical continuous variable) is depicted in (b). An analogous system governed by single electron dynamics (charge is a discrete variable) is depicted in (c). Dynamics of networks with quantum links (in which charge transport is phase-coherent) is depicted in (d).

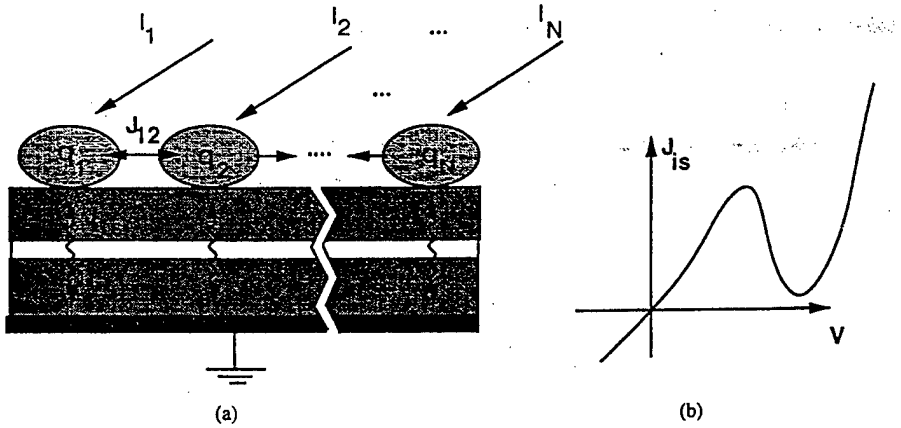
The choice of the system in Fig. 1(a) was motivated by the observation that complex spatial and temporal patterns of the electronic charge distributions on the metallic islands, resulting from the nonlinear interactions amongst thousands of devices (islands) which are collectively driven far from equilibrium by currents applied at the program nodes, may reveal a fundamental kind of computational effect [22]. Indeed, we will show in Section III that this system is tailor made for neuromorphic computation and signal processing.

## III. Specific Examples of Neuromorphic Networks Using Quantum Dot Arrays

In this section, we will provide specific examples of useful computational or signal processing functions performed by the generic array of electrically linked nanometer sized metallic islands shown in Fig. 1(a). We will consider, primarily, memory-like circuits which operate on the basis of neuromorphic principles. It is possible to establish a connection between neuromorphic computational models and the charge dynamics in an array of metallic islands, in the presence of each of the different network mechanisms indicated in Fig. 1(b)–(d). For example, when the operation of the network is governed by single-electron effects at each island [see Fig. 1(c)], the choice of a neuromorphic approach is motivated by a straightforward comparison between the granular dynamics of electrons in an array of islands, and the dynamics of discrete Hopfield networks [23]. We will illustrate this in Section III-C with a specific example. Additional impetus for restricting attention to memory applications comes from the realization that single-electron charging dynamics is stochastic in nature, which hinders the conceptualization of logic applications. Associative memory, on the other hand, can benefit from an appropriate kind of stochasticity.

### A. Neuromorphic Continuous Charge Network

The first example of neuromorphic networks we consider is a purely classical example where the islands in Fig. 1(a) are large enough (diameter ~100 nm) and the temperature high enough that single electron effects are not prominent. Charge can be considered a continuous (rather than discrete) variable and the corresponding network mechanism is illustrated in Fig. 1(b). A one-dimensional (1D) version of the circuit is shown in Fig. 2 in which we assume that the current between islands $i$ and $j$ is $J_{ij}$, and that the current between island $i$




(a) (b)

Fig. 2. (a) A 1D array of islands, whose dynamics is described by (2); (b) a nonmonotone substrate nonlinearity of the kind shown here, is the minimal condition for the realization of collective effects. This characteristic can be produced by a resonant tunneling diode.

and the grounded substrate is $J_{is}$. Since charge is considered a continuous variable, we can write from Kirchoff's current balance condition that

$$\frac{d}{dt}q_i = C_i\frac{d}{dt}v_i = -\left[J_{is}(v_i) + \sum_{j\neq i} J_{ij}(v_i - v_j)\right] + I_i(t) \qquad (1)$$

where $I_i(t)$ is the driving current, $q_i$ is the charge, $v_i$ is the potential, and $C_i$ is the capacitance, with the subscript $i$ indicating the relevant island. $J_{is}$ is a nonmonotonic nonlinear function of $v_i$ corresponding to the current-voltage characteristic of an RTD.

*1) Operation as an Associative Memory:* Qualitatively, the multiple solutions to the nonlinear system of equations which results from (1) under steady-state conditions ($dv_i/dt = 0$ for all $i$) will be taken as a set of memory states which can be programmed by properly choosing the current biases $I_i$. The current biases $I_i$ will be assumed to be either time independent, or slowly varying on time scales over which the network relaxes into its memory states. If the network is begun at time $t = 0$, with a certain initial condition $v_i = q_i/C_i$ for all $i$, arranged, for instance, by an initial impulse of charges $q_i$ at each of the islands, then the network will evolve toward the closest memory state, as programmed by the current biases $I_i$. This is the functioning of an associative memory.

Equation (1) bears strong similarities to the equations representing the additive short-term memory (STM) (or Content Addressable Memory (CAM)) model of neural networks [24]. Therefore, it is obvious by analogy that the system in Fig. 2 realizes the STM or CAM model. The significant difference is that, while conventional circuit realizations of the STM model require operational amplifiers (whose transfer functions exhibit monotonic sigmoidal nonlinearities) and massively interconnected networks of resistors [24]–[26], all we need is an RTD as a substrate which provides a *nonmonotonic* nonlinearity in the local transport. In [8], we showed that this nonmonotonic nonlinearity is sufficient for the realization of associative memory effects. Op-amps are very costly in terms of real estate on a chip and also consume too much power. As a result, neural networks have been unable to compete with conventional silicon VLSI technology. In contrast, RTD's are very compatible with low-power nanoelectronics. They can be integrated vertically beneath the array of islands so that they do not use up any real estate on the chip. There has been some prior research involving resonant tunneling diodes (RTD) for the realization of neuromorphic systems by Levy and co-workers [27]. Additionally, some early investigation of networks of tunnel diodes has been reported by Stern [28] and by Wilson [29], which also exhibit current-voltage characteristics of the kind shown in Fig. 2(b).

In [8], we proved two important results pertinent to this system. First, based on global stability analysis using Liapunov functions, we showed that the system in (1) is globally stable. This implies that there are no limit cycles, and that every trajectory of the system will converge to one of its equilibrium points. Second, we showed that nonmonotonic nonlinearities in the current-voltage characteristic of local transport, such as that shown in Fig. 2(b), are necessary to obtain *multiple* equilibrium points for nontrivial collective computational activity in these networks. It should be contrasted with the general requirement of monotonic nonlinearities in generic additive short-term memory systems [24]–[26] which is realized through op-amps. This result also has immediate consequences to nanoelectronics, in that staircase nonlinearities, which have been produced in quantum point contact constriction [30] and asymmetric double junction single electron tunnel devices [31] can be seen to be inadequate for the realization of collective activity. However, there are several electronic devices which can produce nonmonotonic nonlinearities, e.g., Esaki diode, IMPATT diode, BARITT diode, etc. [32]. We are interested chiefly in semiconductor heterostructure devices which can be integrated vertically beneath the array of islands. They must also be low power devices. Therefore, the ideal structure is a RTD. Since our work focuses on nanoelectronic realizations, complex circuitry cannot be integrated with the basic elements anyway owing to the small size of the islands. Therefore, an RTD seems to be the optimum choice.

We next provide numerical results based on particular choices of $J_{ij}$'s and $J_{is}$'s. Our first example is illustrated in Fig. 3(a), in which the two islands are coupled with each other

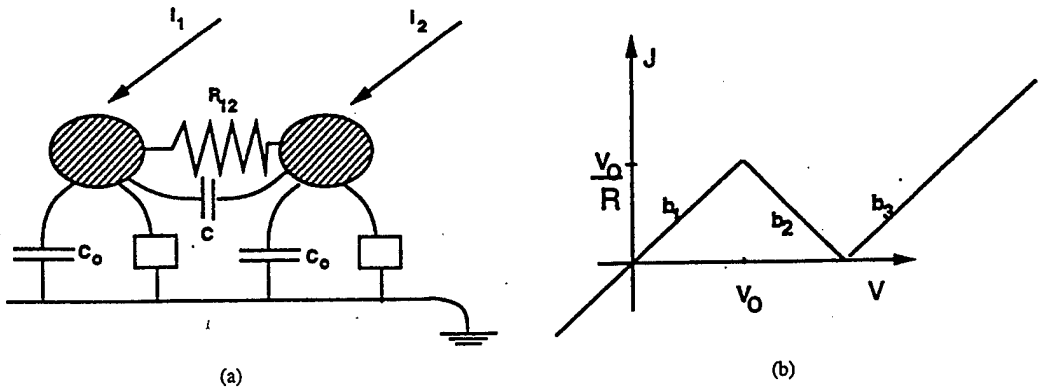



Fig. 3. (a) Two islands coupled by a resistor $R_{12}$ and capacitor $C$. Each island is also coupled to the substrate by a capacitance $C_0$ and a nonlinear resistor whose current voltage characteristic is given in (b); (b) Piecewise approximation to the nonmonotonic nonlinearity of the substrate conduction [cf. Fig. 2(b)]. This characteristic defines the quantities $R$ and $v_0$. Each segment has the same magnitude of slope $1/R$.

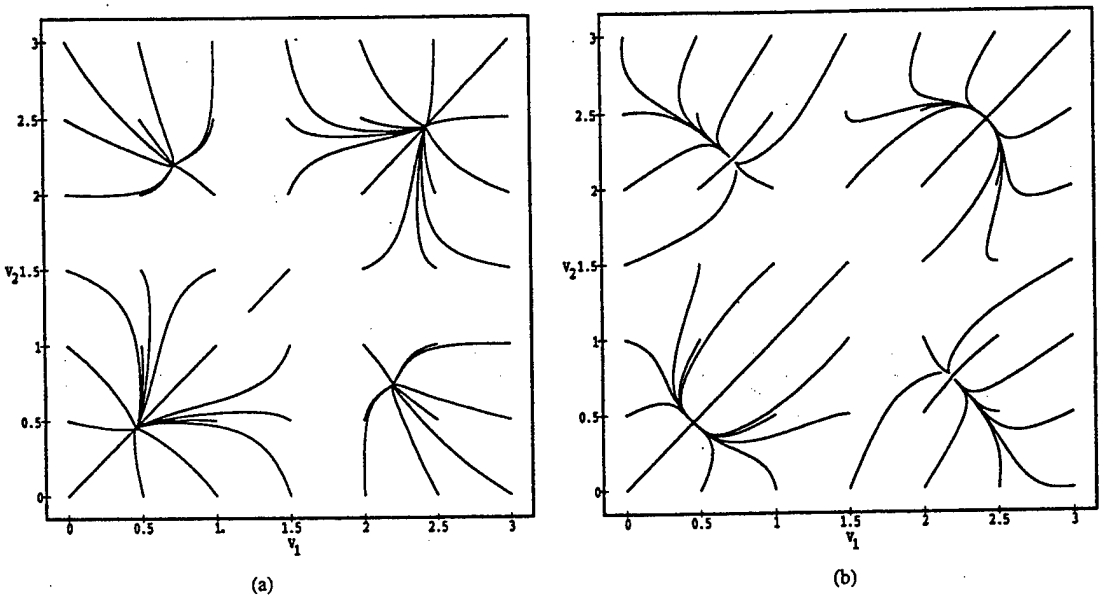



Fig. 4. Phase portrait for two-island system in the continuous-charge model, in which the voltage axes have been scaled with respect to $v_0$. The parameters are $R_{12} = 5R$, and $I_0 = v_0/2R$. (a) with only substrate capacitance $C_0 = 2q/v_0$., and (b) with identical mutual and substrate capacitances: $C = C_0$.

through a linear resistor $R_{12}$ and a capacitor $C$. In addition, we assume that each island is *individually* coupled to the substrate through a capacitance $C_0$, and a nonlinear resistor whose current-voltage characteristic is shown in Fig. 3(b). Furthermore, each island is driven by a current $I_o$ fed externally. Since the results presented here are numerical in nature, we have minimized the number of parameters by choosing a piecewise linear function to represent the I-V characteristic of the RTD (substrate conduction). Each of the three segments are given slopes of the same magnitude $R$

$$J_s(v) = \begin{cases} v/R & v < v_0 \\ (2v_0 - v)/R & v_0 \le v \le 2v_0 \\ (v - 2v_0)/R & v > 2v_0 \end{cases}. \qquad (2)$$

The equilibrium points of this two-node system can be found by setting $\dot{v}_1 = \dot{v}_2 = 0$, and the stability properties of those equilibrium points can be ascertained by examining the eigenvalues of the system matrix near the equilibrium points [33], [34]. Provided that the driving current $I_0 < v_0/R$, each island can in principle be on any one of the three branches of the nonlinear function $J_s(v)$. We can, however, show that the system will be unstable if either island is operated on branch $b_2$. So, it follows that this system can have at most four different globally stable points, since each island is restricted to being on either branch $b_1$ or $b_3$. Two of these stable points are trivial ones corresponding to both islands operating on the same branch : either $b_1$, or $b_3$. We have determined a necessary and sufficient condition for the existence of all four "memories," and the corresponding phase portraits are shown in Fig. 4.

One can also realize *programmed* associative memory. An example is shown in Fig. 5 where different choices of inter-island resistances lead to the occurrence of either eight, four, or two equilibrium points.

*2) Image Processing Network:* Fig. 6 shows results of further numerical investigations, where a 2D array of 10 × 10 islands with *only near-neighbor connections* and the same substrate nonlinearity [as shown in Fig. 3(b)] are used. An

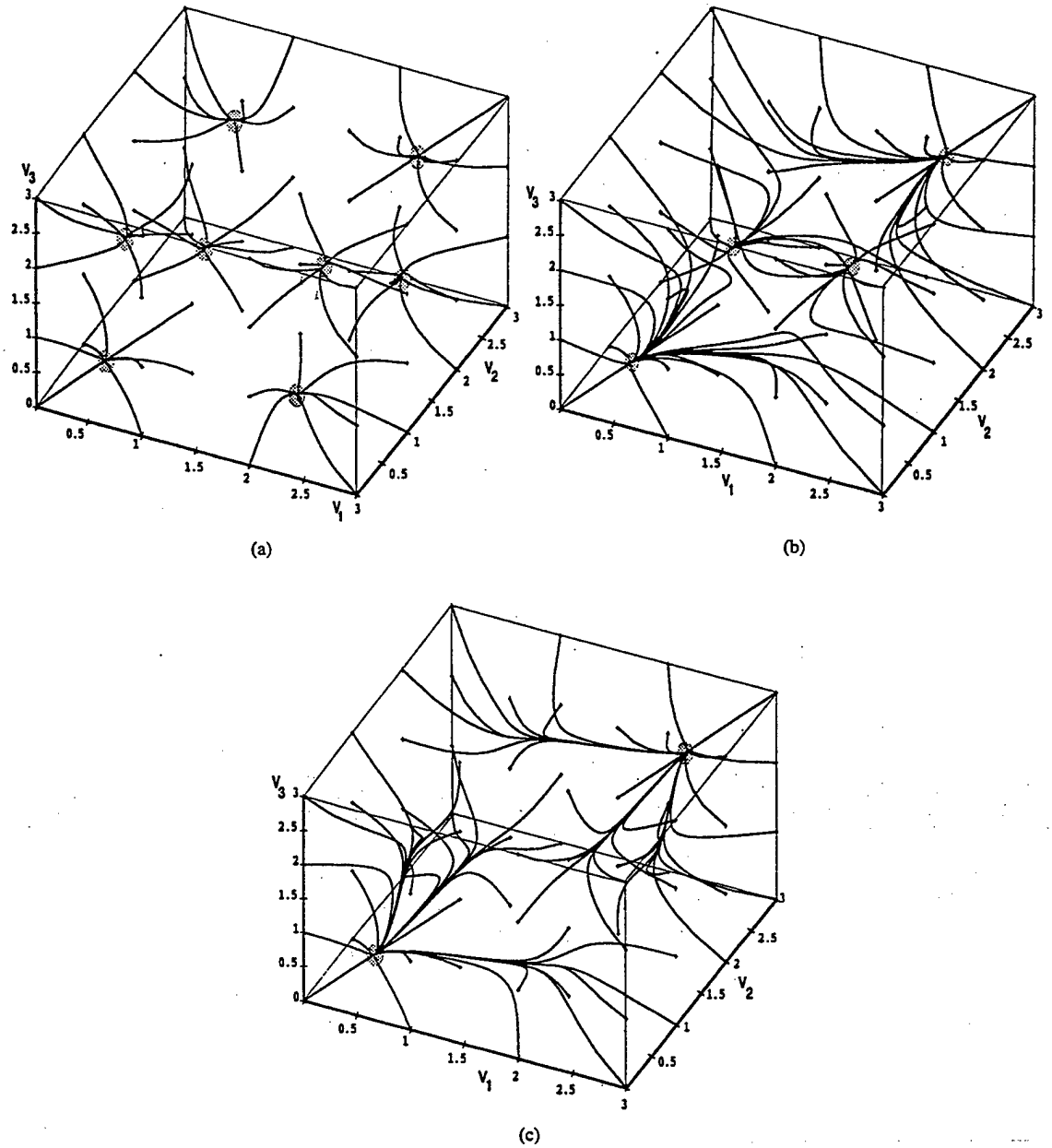



Fig. 5. Phase portrait for a three island system. This figure indicates the possibility of programming a network by choosing resistance parameters appropriately. Only substrate coupling $C_0$ is permitted. The stable equilibrium points are indicated by the shaded circles. (a) eight states, with $R_{12} = R_{23} = R_{13} = 10R$.; (b) four states, with $R_{12} = R_{23} = 10R$. and $R_{13} = 5R$.; (c) two states, with $R_{12} = R_{23} = R_{13} = 5R$.

input image is digitized to create dark and light pixels. Each pixel generates a voltage on a corresponding island. These voltages lie on a particular branch—$b_1, b_2$, or $b_3$, in Fig. 3(b)—of the substrate nonlinearity. Dark pixels place the corresponding voltages on the branch $b_3$, light pixels on the branch $b_1$ and intermediate grey pixels on the branch $b_2$. In example (a), the input contains domains which are either predominantly dark, or light, and the resulting output recovers domains which are either all dark, or all light. This is an example of "smoothing," an important image processing capability. Other image processing functions can be realized as well. We must point out that this 10 × 10 network consumes an area of $10^{-10}$ cm$^2$ so that ten trillion such image processors can be fabricated within an area of 1 cm$^2$. This is an unprecedented functional density.

### B. *Single-Electron Charging Effects*

We have also considered the same two-island network as in Section III-A, with the assumption that single electron charge effects (granularity of charge) have become pertinent either due to the lowering of temperature, or due to the physical scaling of the metallic islands down to $d \sim 10$ nm. Single-electron effects will become relevant, when the change in potential $\delta V = q/C$, associated with the addition of a single electron charge $q$ to an island, becomes comparable to $kT/q$, the thermal potential. This condition can be met even at room

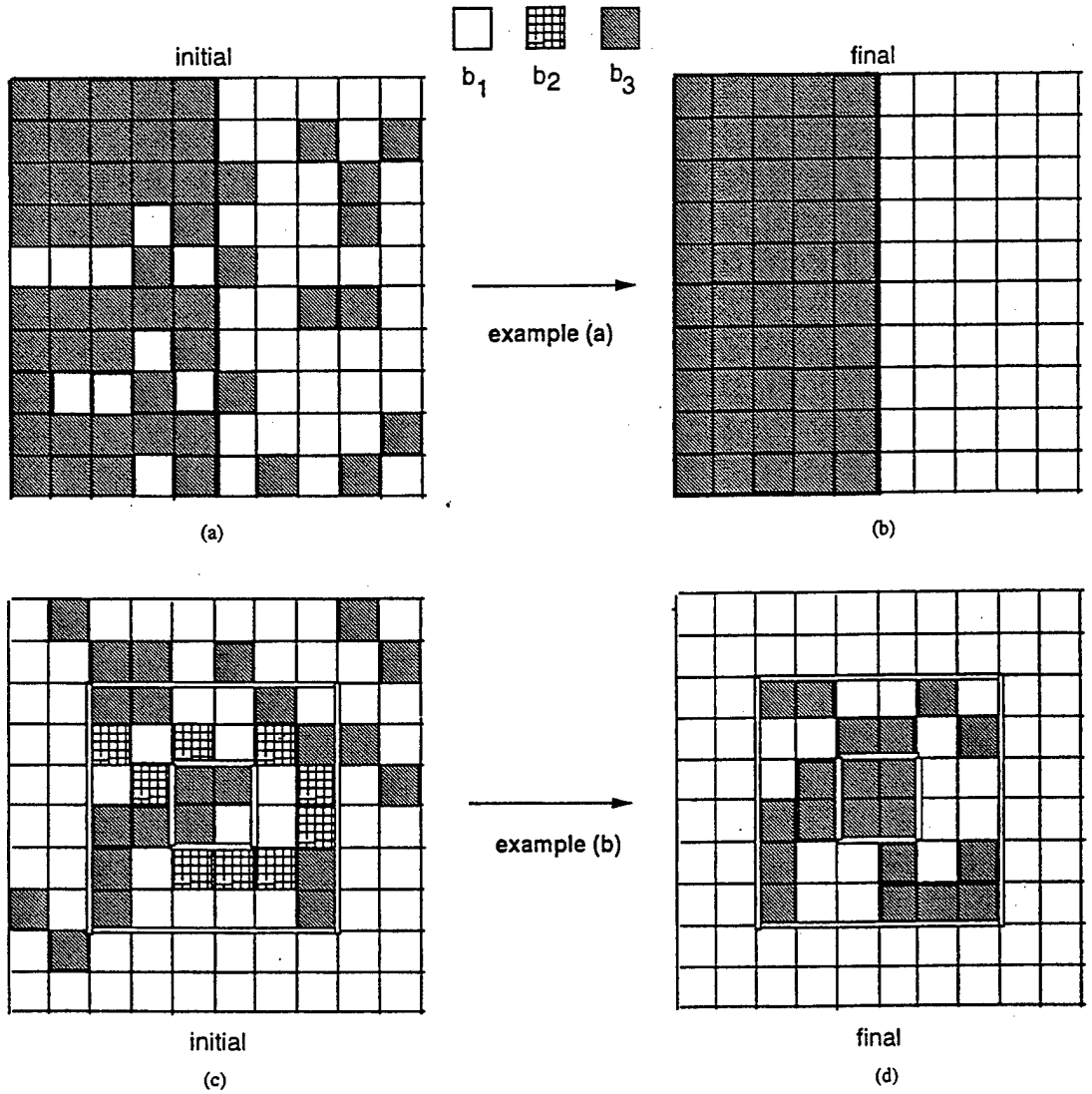



Fig. 6. Rudimentary image processing capability in a *near-neighbor* connected network of 10 × 10 islands, which are all pumped by the same current $I_0 = v_0/2R$. Each island is colored in accordance with the particular branch—$b_1$, $b_2$, or $b_3$, in Fig. 3(b)—of the substrate nonlinearity, the island potential lies on. (a) The resistive network is uniform $R_{i\pm1,j\pm1} = 6R$. The input contains domains which are either predominantly black, or white, and the resulting output recovers domains which are either all black, or all white. (b) The network is partitioned into three concentric regions. The islands in the outermost and innermost regions are coupled with low resistances $R_{i\pm1,j\pm1} = 3R$, and the islands in the intermediate region are coupled with high resistances $R_{i\pm1,j\pm1} = 13R$. Parts of the network with low resistance produce regions which are either all white, or all black depending on which was predominant in the initial state.

temperature if islands with an effective capacitance smaller than $C \approx 5 \times 10^{-18}$ F are fabricated.

Single electron dynamics are characterized by discrete tunnel events. Therefore, the differential equation systems of (2) cannot be used to model the arrays. We developed a Monte-Carlo simulation technique [35], [36], for the simulation of a current biased network of islands, shown in Figs. 1 and 2. Owing to space limitations, we cannot describe it here, but instead refer the reader to [8].

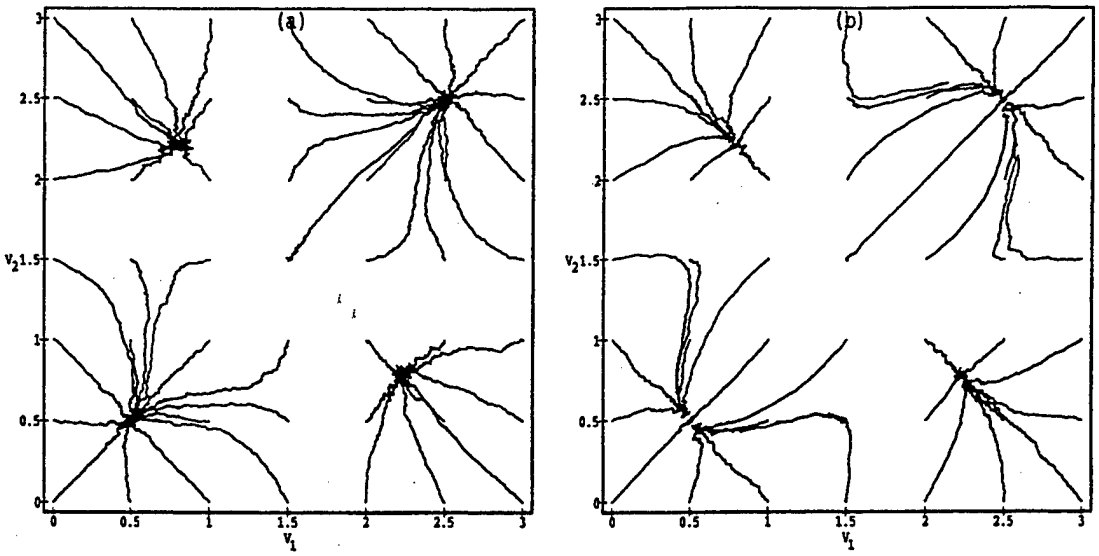
The results based on our numerical simulator show the following:

1) For large enough values of the effective substrate capacitances ($C_0$), the effect of the single electron dynamics is marginal and the *system is adequately described by the continuous charge models*. Thus even arrays of metallic islands with diameters of few nanometers can exhibit the same associative memory and image processing capabilities studied in Section III-A.
2) If the substrate capacitance is made smaller, the granularity of the trajectories become larger, and for small enough values, the phase portrait breaks out into oscillations [8].

These effects are illustrated in Fig. 7, where the same system shown in Fig. 3 is investigated under single-electron dynamics.

### C. Ising-Type Single-Electron Networks

The examples in Section III-A and III-B did not utilize single-electronics in a direct way, i.e., single electron charging effect was not central to the operation of the circuit. However, there are instances when single electronics plays a pivotal role. In [8], it was shown that the evolution of an initial charge distribution toward a stable final equilibrium distribution in a




(a) (b)

Fig. 7. Phase portrait for the same two island system in Fig. 3 where the granularity of charge (single electron charging) has been taken into account. These results are obtained from Monte-Carlo simulation and are discrete charge versions of the two cases in—Fig. 4(a).

system of the type of Fig. 1(a) can be given a neuromorphic interpretation. These properties emerge *purely as a result of the discreteness of electronic charge*, and the only assumption made in establishing this connection is that the inter-island capacitances can be chosen arbitrarily.

In [8], it was shown that the discrete stochastic dynamics of electrons can be viewed as a procedure for simulated annealing. Consequently, several problems from combinatorial optimization can then be mapped to different single-electron networks. A novel mapping methodology tailored specifically for single-electron networks was developed. Owing to space limitations, we cannot address this in greater detail here, but we mention that a specific example we examined is the traveling salesman problem. This NP-complete optimization problem can be solved very fast by an appropriate single electron network. For details, we refer the reader to [8].

## IV. Fabrication of Neuromorphic Networks: Self-Assembly

We now address the fabrication of these networks. Conventional nanolithography involving electron beam, ion beam, x-ray or even scanning tip lithography [37] are unsuitable for fabricating single-electron circuits for two basic reasons: 1) they are usually *serial* in nature (direct-write techniques) whereby the patterns on each wafer must be exposed sequentially, resulting in a throughput that is unacceptably slow for commercial mass production; and 2) they often cause significant radiation damage to processed nanostructures [38]–[43]. In contrast, the recently developed techniques for self-assembly of periodic quantum dot arrays [14], [44] suffer from neither of these drawbacks and therefore offer an attractive choice.

Fabrication of the network shown in Fig. 1(a) has three key elements: 1) deposition of a regimented, *periodic* array of nanosized metallic islands on a RTD; 2) electrical isolation of these islands via mesa-etching through the RTD; and 3) linking the islands with molecular wires or some other means. Fabrication of a periodic, 2D array of metallic islands with diameters in the 5–100 nm range has been demonstrated by us using electrochemical self-assembly [44]. This approach has two advantages over conventional direct-write lithography (e.g., e-beam, ion-beam or STM lithography). First, it is a "gentle" technique that does not cause serious processing damage, and second, it is *parallel* in nature whereby large batches of wafers can be processed simultaneously and all patterning on a given wafer is performed simultaneously, resulting in several orders of magnitude improvement in throughput. In order to illustrate the feasibility of realizing nanoscale circuits using these methods, we outline a specific process which employs self-assembly techniques, several of which have been developed by the authors and their collaborators.

First, using molecular beam epitaxy, an RTD structure is synthesized and capped by a thin (3 nm) low-temperature grown GaAs layer (LTG:GaAs), which is highly resistant to oxidation [45] and which serves to passivate underlaying doped layers [45], [46]. Then on the surface of this layer, a thin-film of Al (~100-nm-thick) is resistively evaporated. This film is electropolished in perchloric acid/butyl cellusolve/ethanol/distilled water at 60 V for 30 s resulting in a dimpled surface containing a highly periodic array of crests and troughs as shown in the *raw* atomic force micrograph data in Fig. 8.

The trough regions of this dimpled film can be selectively etched away by bromine/methanol leaving isolated islands of Al (the crests) on the surface (see Fig. 9). The trough regions are etched away before the crests since the Al film is much thinner in these regions (by approximately 50 nm). In fact, the ratio of the Al film thicknesses in the crest and trough regions can be easily made as large as 10:1 which facilitates selective etching of the troughs. This results in a 2D periodic array of nano-sized metallic islands on the surface (vestiges of the crests) which now need to be electrically isolated. For this, we

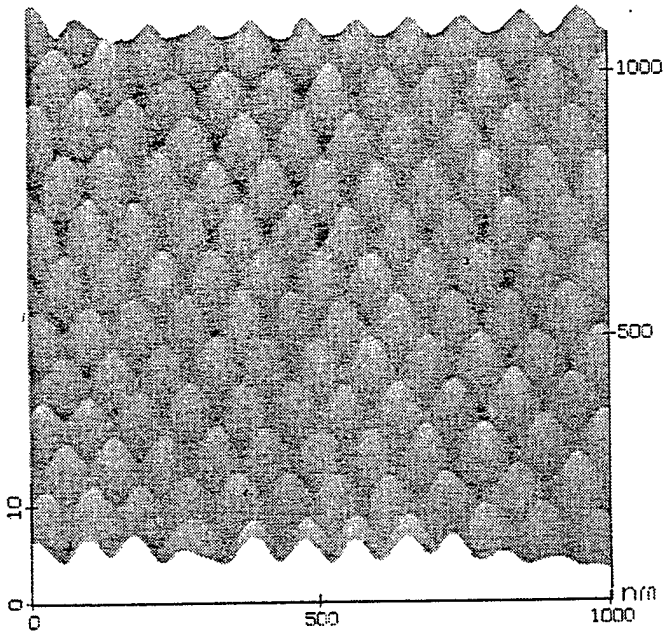



Fig. 8. Atomic force micrograph of a dimpled electropolished Al surface showing a periodic array of crests and troughs. This pattern is self-organized. The height of the crests above the troughs is about 50 nm.

proceed as follows. One can use this periodic array of islands as a mask to shallow-etch electrically isolated mesas capped by the metal Al. The use of the metallic island array as the etching mask for patterning the RTD substrate provides a self-alignment between the metal islands and the semiconductor mesas. Deep etching is not feasible, due to the small areas of the Al dots and the need to keep a relatively planar surface for subsequent interconnect steps. The LTG:GaAs layer allows one to keep the underlying n-GaAs layer very thin without risking oxidation [45]. The isolation etch can be shallow since the removal of the LTG:GaAs layer eliminates the passivation effect, resulting in the depletion of the regions between the islands. The final step is to electrically link the islands. For this, the space between the mesas may be filled up with a chosen material by electrodeposition, sputtering, or low-temperature grown MBE layer to create a resistive and/or capacitive connection between the dots.

Instead of using the electropolished Al technique, one can also deposit a self-assembled, 2D network of metal clusters on the surface of the substrate with adjacent clusters linked by molecular wires. Techniques to deposit ordered arrays of nanometer diameter Au clusters on various flat substrates, to deposit organic molecules (molecular wires) between adjacent clusters have been demonstrated and the in-plane electronic conduction through arrays of 4 nm clusters deposited between lithographically defined contacts has been studied [14], [15], [47]. Again, the deposited cluster network can be used as a self-aligned mask to etch through the RTD and create mesas capped by metal dots (Fig. 10). The inter-island links are already provided by the molecular wires. Although the technology of molecular wires is in its infancy, significant progress has been made in the synthesis, characterization and modeling of molecular wires suitable for this application [48], [49]. The advantages of molecular wires are: 1) great flexibility in choosing the resistance and capacitance between clusters [48]; and 2) electrical access to individual nanometer sized dots may be feasible due to the nanometer size scale of the wires and the ability to deposit wires on selective regions based on the composition of the regions and of the molecular end groups. These capabilities would be necessary for realizing certain processors where one needs variable inter-island resistances and capacitances as well as image processors where one needs to address individual dots.

The fabrication steps outlined to this point have addressed the metallic island array, the interisland linking and the RTD substrate required for the computational cells. The same techniques used to provide resistive linking between islands can provide the linking between islands at the edge of the array and input/output or programming contacts. In "edge-driven" configurations, the computational cells will also require a resistive link between each island and a bias plane, presumably located above the island array, to provide the current bias required to achieve bistability in voltage. This capability appears feasible using either molecular wires or a thin-film resistive layer deposited between the island array and the bias plane. It should be noted that the most time consuming "direct-writing" steps of nanolithography have been replaced by chemical self-assembly in the proposed synthesis techniques. In principle, large batches of wafers can be electropolished and subsequently etched in a large bath, thereby making the electropolishing technique a truly parallel process with fast throughput. The self-assembly synthesis of metallic cluster arrays could also be a high throughput process. This is a significant advantage over "direct" write based fabrication processes for nanoscale circuits.

The ability to realize individual address lines for each island, while beyond the capability of the self-assembly techniques presented here, would allow configurations where individual

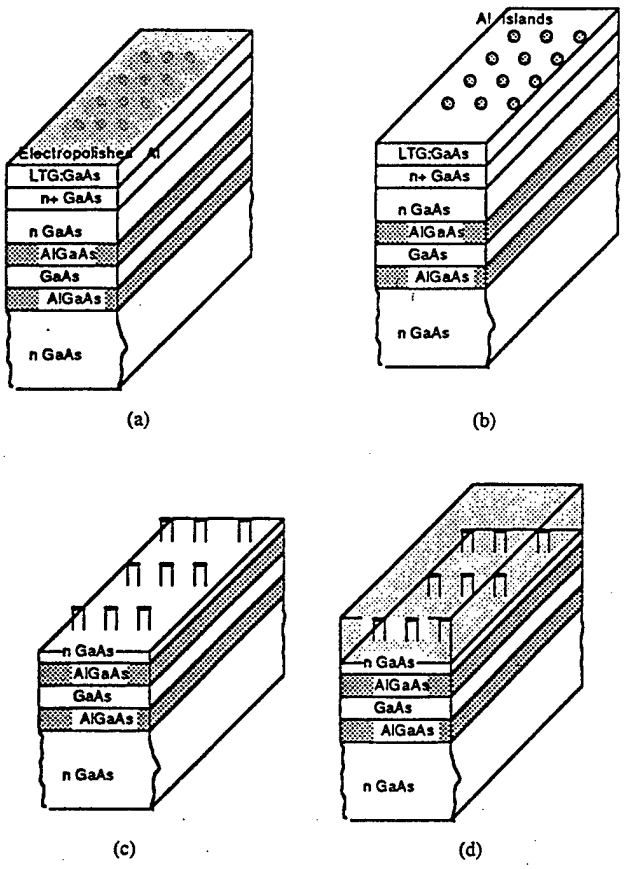



Fig. 9. Fabrication steps to realize the neuromorphic network in Fig. 1(a) using self-assembly techniques. The steps are described in the text. The inter-island connection is made *not* by molecular wires, but by an electrodeposited, sputtered or MBE-grown layer grown on the etched surface. A wide variety of layer materials can be used, resulting in phase-coherent, rectifying, resistive or capacitive links.

islands could be addressed in order to provide input/output or programming, Achieving this type of interconnect configuration at the nanometer scale would require significant advances in technology, but may be possible using scanning probe based lithography.

In order to establish the feasibility of the proposed synthesis techniques, it is informative to consider the ranges of capacitance and resistances achievable with these techniques. Using a conservative estimate for the minimum achievable island diameter of 20 nm, the capacitances between adjacent islands and between an island and the substrate will be on the order of $10^{-18}$ F, corresponding to Coulomb charging energies of about 70 meV. Molecular wires which have been reported to date have resistances on the order of 20–40 M$\Omega$ per molecule when bridged between metal electrodes [15], [48], [49]. Since metallic clusters with 20-nm diameters can have hundreds of molecular wires in parallel between each pair of adjacent clusters, achievable island to island resistances should be on the order of 0.1–10 M$\Omega$ [15], [47], [49]. In order to bridge distances significantly longer than 2 nm, networks of these molecular wires may be required. An RTD substrate with peak current density of $1 \times 10^5$ A/cm$^2$ would yield a peak current of 0.3 $\mu$A for each semiconductor mesa, assuming that the current density remained constant as device dimensions

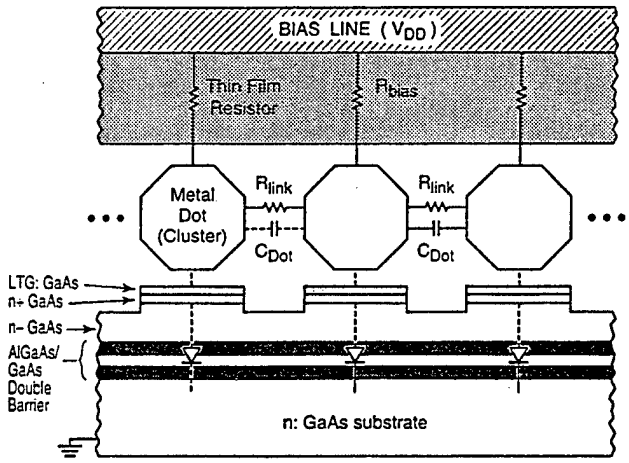



Fig. 10. Another possible fabrication strategy where the metal dots are clusters self-assembled on the surface and linked by molecular wires. $R_{link}$ and $C_{Dot}$ are the interdot resistances and capacitances that are realized by molecular wires. Self-assembly techniques for 2D linked arrays have been demonstrated and the in-plane electronic conduction through such an arrays has been studied (see [14] and [48]).

are scaled down. For a voltage at peak current of 0.5 V, this corresponds to an average resistance of approximately 1.5 M$\Omega$ for the RTD characteristic. The use of short-chain molecular wires along with a low-resistance, nonalloyed contact structure [50] should provide an interface coupling resistance of less than 200 K$\Omega$ between the metal island and the semiconductor mesa for this island size. The devices can be effectively current biased if the resistance from a bias plane, located above the array, to each island is much greater than the characteristic resistance of the RTD mesa. Since the resistance through certain molecular wires consists of state-assisted tunneling [48], a suitable resistance, e.g., a resistance on the order of 5–10 M$\Omega$, could be achieved by using longer molecular wires to connect each island to the bias plane. It may also be possible to use a thin-film resistive material deposited between the metallic island array and the bias plane to realize the linking of each island both to adjacent islands and to the bias plane. A resistivity of approximately 50 $\Omega \cdot$ cm would be required; an ideal material would also possess an anisotropic resistivity characteristic to allow independent control of the vertical and horizontal linking resistances. It therefore appears feasible to realize appropriate resistance and capacitance elements for the proposed computational cells using the technologies discussed in this section.

## V. Conclusion

In conclusion, we have presented a neuromorphic architecture for powerful computing and signal processing applications using an array of metallic islands on a resonant tunneling diode. Some of the components required to fabricate these circuits are in their infancy. Nontheless, the rapid progress in nanoelectronics may make these circuits a reality in not too distant a future.

## Acknowledgment

The authors are indebted to numerous colleagues for insightful discussions. We are especially grateful to Dr. J. A. Cooper,

Jr., Dr. P. Balasingam, Dr. M. Cahay, Dr. A. E. Miller, Dr. R. P. Andres, Dr. J. M. Woodall, and Dr. R. Landauer.

## REFERENCES


[1] R. T. Bate, G. A. Frazier, W. R. Frensley, J. W. Lee, and M. A. Reed, *Proc. SPIE*, vol. 792, pp. 26–35, 1987; R. T. Bate, *Solid State Technol.*, vol. 32, p. 101, 1989.

[2] P. Bakshi, D. Broido, and K. Kempa, *J. Appl. Phys.*, vol. 70, p. 5150, 1991.

[3] H. Korner and G. Mahler, *Phys. Rev. B*, vol. 48, p. 2335, 1993; K. Obermayer, W. G. Teich and G. Mahler, *Phys. Rev. B.*, vol. 37, 1987, 8096; W. G. Teich and G. Mahler, *Phys. Rev. A.*, vol. 45, p. 3300, 1992.

[4] S. Lloyd, *Science*, vol. 261, p. 1569, 1993; S. Lloyd, *Science*, vol. 263, p. 695, 1994.

[5] C. S. Lent, P. D. Tougaw, W. Porod, and G. H. Bernstein, *Nanotechnology*, vol. 4, p. 49, 1993.

[6] S. Bandyopadhyay, B. Das, and A. E. Miller, *Nanotechnology*, vol. 5, p. 113, 1994.

[7] A. N. Korotkov, *Appl. Phys. Lett.*, 67, 2412 1995.

[8] P. Balasingam and V. P. Roychowdhury, "nanoelectronic Functional Devices," School of Electrical Engineering, Purdue Univ., West Lafayette, IN, Tech. Rep. TR-EE 94-24, 1994.

[9] S. Bandyopadhyay and V. P. Roychowdhury, *IEEE Potentials*, p. 8, Apr.-May 1996; S. Bandyopadhyay and V. P. Roychowdhury, *Jpn. J. Appl. Phys.*, pt. 1, vol. 35, p. 3350, 1996; V. P. Roychowdhury, X. Wang, and S. Bandyopadhyay, in presented at *8th Int. Conf. Superlattices, Microstructures, and Microdevices*, Cincinnati, OH, Aug. 24, 1995.

[10] R. A. Kiehl and T. Ohshima, in *Extended Abstracts of the International Conference on Solid State Devices and Materials, SSDM'95*, Osaka, Japan, p. 210. To appear in *Jpn. J. Appl. Phys.*.

[11] R. Landauer, keynote talk given at the *Workshop on Physics and Computation, PhysComp '92*, Dallas, TX, Oct. 2-4, 1992, (available as an IEEE Computer Society Press Reprint); R. Landauer, in *"Ultimate Limits of Fabrication and Measurement"*, (NATO-ASI Series), M. E. Welland, ed. Dordrecht: Kluwer, 1994.

[12] ———, *IBM J. Res. Develop.*, vol. 5, p. 183, 1961.

[13] R. Tsu and L. Esaki, *Appl. Phys. Lett.*, vol. 22, p. 562, 1973.

[14] R. G. Osifchin, R. P. Andres, J. I. Henderson, and C. P. Kubiak, to be published; R. G. Osifchin, W. Mahoney, R. P. Andres, J. I. Henderson, C. P. Kubiak, and R. N. Dominey, *Superlattices and Microstructures* vol. 18, pp. 283–289, 1995.

[15] D. B. Janes, V. Kolagunta, R. G. Osifchin, and R. P. Andres, "Electronic conduction through two-dimensional arrays of nanometer diameter metal clusters," *Superlattices and Microstructures*, vol. 18, pp. 275–282, 1995.

[16] G. M. Whiteside, *Sci. Amer.*, p. 146, Sept. 1995.

[17] D. S. Bethune, R. D. Johnson, J. R. Salem, M. S. deVries, and C. S. Yannoni, *Nature*, vol. 366, p. 123, 1993; N. Sudhakar, S. V. Sharma, L. S. Grogoryan, P. Chand, A. K. Majumdar, P. K. Dutta, S. B. Samanta, and A. V. Narlikar, *Phys. Lett. A*, vol. 170, p. 235, 1992.

[18] V. Bogomolov, Y. Kumzerov and S. Romanov, *Physics of Nanostructures*, J. H. Davies and A. R. Long, Eds. Edinburgh, UK: Scottish Universities Summer School in Physics Publications, 1992, p. 317.

[19] G. Binnig, H. Rohrer, C. Gerber, and E. Weibel, *Appl. Phys. Lett.*, vol. 40, p. 178, 1982; G. Binnig, H. Rohrer, C. Gerber, and E. Weibel, *Phys. Rev. Lett.*, 49, p. 57, 1982.

[20] M. I. Lutwwyche and Y. Wada, *J. Vac. Sci. Technol. B*, vol. 13, p. 2819, 1995.

[21] E. J. P. Santos and M. C. McDonald, *J. Vac. Sci. Technol. B*, vol. 11, 2362, 1993.

[22] G. Nicolis and I. Prigogine, *Exploring Complexity*. New York: Freeman, 1989.

[23] J. J. Hopfield, in *Proc. Natl. Acad. Sci.*, vol. 79, p. 2554, 1982; J. J. Hopfield and D. W. Tank, *Science*, vol. 233, 625, 1986.

[24] S. Grossberg, *Neural Networks*, vol. 1, p. 17, 1988.

[25] M. A. Cohen and S. Grossberg, *IEEE Trans. Syst. Man, Cybern.*, vol. SMC-13, p. 815, 1983.

[26] A. N. Michel, J. A. Farrel, and W. Porod, *IEEE Trans. Circuits Syst.*, vol. 36, p. 229, 1989.

[27] H. J. Levy, D. A. Collins, and T. C. McGill, in *Proc, IEEE Int. Symp. Circuits and Syst.*, New York, 1992; also in *Tech. Dig. ISCAS*. New York: IEEE Press.

[28] T. E. Stern, *Theory of Non-linear Networks and Systems*. Reading, MA: Addison-Wesley, 1965.

[29] A. N. Wilson, *Non-linear Networks: Theory and Analysis*. New York: IEEE Press, 1974.

[30] B. J. van Wees, *et al.*, *Phys. Rev. Lett.*, vol. 60, p. 848, 1988; D. A. Wharam, *et al.*, *J. Phys. C.*, vol. 21, p. L209, 1988.

[31] M. Amman, R. Wilkins, E. Ben-Jacob, P. D. Maker, and R. C. Jaklevic, *Phys. Rev. B*, vol. 43, p. 1146, 1991.

[32] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. New York: Wiley, 1981.

[33] A. A. Andronov, A. A. Vitt, and S. E. Khaikin, *Theory of Oscillators*. New York: Dover, 1966.

[34] D. W. Jordan and P. Smith, *Non-linear Ordinary Differential Equations*. Oxford, UK: Clarendon, 1987.

[35] N. S. Bakhvalov, G. S. Kazachca, K. K. Likharev, and S. I. Serdyukova, *Sov. Phys. JETP*, vol. 68, p. 581, 1989.

[36] U. Geigenmüller and G. Schön, *Europhys. Lett.*, vol. 10, p. 765, 1989.

[37] E. E. Ehrichs, R. M. Silver, and A. L. deLozanne, *J. Vac. Sci. Technol. A*, vol. 6, p. 540, 1988.

[38] B. Das, S. Subramaniam, and M. R. Melloch, *Semicond. Sci. Technol.*, vol. 8, p. 1347, 1993.

[39] D. C. Tsui, A. C. Gossard, and G. J. Dolan, *Appl. Phys. Lett.*, vol. 42, p. 180, 1983.

[40] T. Fink, D. D. Smith, and W. D. Braddock, *IEEE Trans. Electron Devices*, vol. 37, p. 1422, 1990.

[41] T. B. Stellwag, M. R. Melloch, J. A. Cooper, Jr., S. T. Sheppard, and D. D. Nolte, *J. Appl. Phys.*, vol. 71, p. 4509, 1992.

[42] C. M. Knoedler, *J. Appl. Phys.*, vol. 68, p. 1129, 1990.

[43] M. R. Melloch, private communication.

[44] A. E. Miller, S. Bandyopadhyay, G. Banerjee, D-F Yue, and M. Chandrasekhar, unpublished, 1995; S. Bandyopadhyay, A. E. Miller, and M. Chandrasekhar, in *Proc. of SPIE, Photonics West'95*, San Jose, CA, (PAGE NOS?) Feb., 1995

[45] E. H. Chen, D. McInturff, T. P. Chin, M. R. Melloch, and J.M. Woodall, *Appl. Phys Lett.*, vol. 68, p. 1678, 1996.; S. Hong, D. B. Janes, D. McInturff, R. Reifenberger, and J. M. Woodall, "Stability of a low-temperature grown GaAs surface layer following air exposure using tunneling microscopy," to be published.

[46] D. C. Look, D. C. Walters, C. E. Stutz, K. R. Evans, and J. R. Sizelove, "Mechanisms for GaAs surface passivation by a molecular beam epitaxial cap layer grown at 200 C," *J. Appl. Phys.*, vol. 12, pp. 5981–5984, 1992.

[47] R. P. Andres, J. D. Bielefeld, J. I. Henderson, D. B. Janes, V. R. Kolagunta, C. P. Kubiak, W. J. Mahoney, and R. G. Osifchin, "Synthesis and electrical characterization of a two-dimensional superlattice of molecularly linked metal clusters," to appear in *Science*.

[48] M. P. Samanta, W. Tian, S. Datta, J. I. Henderson, and C. P. Kubiak, "Electronic conduction through organic molecules," *Phys. Rev. B*, vol. 53, p. 7626R, 1996; W. Tian, M. P. Samanta, and S. Datta, to be published.

[49] R. P. Andres, T. Bein, M. Dorogi, S. Feng, J. I. Henderson, C. P. Kubiak, W. Mahoney, R. G. Osifchin, and R. G. Reifenberger, "'Coulomb staircase' single electron tunneling at room temperature in a self-assembled molecular nanostructure," *Science.*, vol. 272, pp. 1323–1324, 1996.

[50] M. P. Patkar, T. P. Chin, J. M. Woodall, M. S. Lundstrom and M. R. Melloch, "Very low resistance nonalloyed ohmic contacts using low-temperature molecular beam epitaxy of GaAs," *Appl. Phys. Lett.*, vol. 66, pp. 1412–1414, 1995.

**Vwani P. Roychowdhury** (M'89) received the B.Tech. degree from the Indian Institute of Technology, Kanpur, and the Ph.D. degree from Stanford University, Stanford, CA, in 1982 and 1989, respectively, both in electrical engineering.

From September 1989 to August 1991, he was a Research Associate, Department of Electrical Engineering, Stanford University. Currently, he is an Associate Professor, School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN. His research interests include parallel algorithms and architectures, design and analysis of neural networks, application of computational principles to nanoelectronics, special purpose computing arrays, VLSI design and fault-tolerant computation. He has coauthored several books including *Discrete Neural Computation: A Theoretical Foundation* (Prentice Hall, 1995) and *Theoretical Advances in Neural Computation and Learning* (Kluwer, 1994).

Dr. Roychowdhury was named the General Motors Faculty Fellow in the School of Engineering at Purdue University from 1992 to 1994. For his outstanding performance in the classroom, he received the 1993 Ruth and Joel Spira Outstanding Teacher Award.

**David B. Janes** (S'86–M'89) was born in Somerset, KY, in 1958. He received the B.A. degree in physics (summa cum laude) from Augustana College, Rock Island, IL, in 1980, and the B.S.E.E. (highest honors), M.S.E.E., and Ph.D. degrees from the University of Illinois, Urbana-Champaign, in 1980, 1981, and 1989, respectively. His doctoral research focused on the effect of deep levels on the performance of acoustic charge transport devices.

From 1981 to 1985, he was a Research Scientist in the Research Division, Raytheon Company, where he was involved in research on microwave semiconductor devices, including GaAs IMPATT diodes and MESFET's, and monolithic microwave integrated circuits. Since 1989, he has been an Assistant Professor, School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN. He is presently engaged in experimental research on mesoscopic semiconductor devices and metallic cluster arrays, characterization of novel semiconductor heterostructures, and microwave devices based on compound semiconductor heterostructures.

**Supriyo Bandyopadhyay** (SM'89) received the B. Tech. degree from the Indian Institute of Technology, Kharagpur, in electronics and electrical communications engineering in 1980, the M.S. degree in electrical engineering from Southern Illinois University, Edwardsville, in 1982, and the Ph.D. degree in electrical engineering from Purdue University, West Layfayette, IN, in 1985.

Currently, he is an Associate Professor, Department of Electrical Engineering, University of Notre Dame, South Bend, IN. He has authored about 150 articles in his various fields of interest which include quantum devices, hot electron transport, superlattice physics, nonlinear optics, and superconductivity. He has chaired many international conferences and edited books in these fields.

Dr. Bandyopadhyay is a member of Tau Beta Pi, Eta Kappa Nu, Sigma Pi Sigma, and the American Physical Society. He is also listed in *Who's Who in Science and Technology*, the *Dictionary of International Biography*, (Cambridge, UK), and *Sterling's Who's Who*.

**Xiaodong Wang** received the B.S. degree in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 1992, and the M.S. degree in electrical engineering from Purdue University, West Lafayette, IN, in 1995. He is currently a Ph.D student in the Department of Electrical Engineering, Princeton University, Princeton, NJ. He has worked in the areas of digital communications, and parallel and distributed computing.

# Nanoelectronic Architecture for Boolean Logic

VWANI P. ROYCHOWDHURY, MEMBER, IEEE, DAVID B. JANES, MEMBER, IEEE, AND SUPRIYO BANDYOPADHYAY, SENIOR MEMBER, IEEE

*A nanoelectronic implementation of Boolean logic circuits is described where logic functionality is realized through charge interactions between metallic dots self-assembled on the surface of a double-barrier resonant tunneling diode (RTD) structure. The primitive computational cell in this architecture consists of a number of dots with nearest neighbor (resistive) interconnections. Specific logic functionality is provided by appropriate rectifying connections between cells. We show how basic logic gates, leading to combinational and sequential circuits, can be realized in this architecture. Additionally, architectural issues including directionality, fault tolerance, and power dissipation are discussed. Estimates based on the current–voltage characteristics of RTD's and the capacitance and resistance values of the interdot connections indicate that static power dissipation as small as 0.1 nW/gate and switching delay as small as a few picoseconds can be expected. We also present a strategy for fabricating/synthesizing such systems using chemical self-organizing/self-assembly phenomena. The proposed synthesis procedure utilizes several chemical self-assembly techniques which have been demonstrated recently, including self-assembly of uniform arrays of close-packed metallic dots with nanometer diameters, controlled resistive linking of nearest neighbor dots with conjugated organic molecules and organic rectifiers.*

## I. INTRODUCTION

It is generally accepted in the solid-state community that conventional strategies for integrating electronic devices on a chip will not be suitable for nanometer-sized elements because of the minuscule size of the elements and the corresponding low power handling capacity, low gain, and low fanout. Accordingly, a number of proposals have appeared in the literature that envision novel architectures for nanoelectronic logic circuits. Most of these schemes purport to exploit discrete (single- or few-electron) charge interactions between semiconductor or metallic dots [1]–[10]. Unfortunately, they all suffer from one or more drawbacks. References [3] and [4] propose a paradigm which is plagued by problems associated with unavoidable and ubiquitous background charge fluctuations that make the circuits error prone, [5] requires precise clocking control and lacks sufficient fault tolerance, while [7]–[9] proposes a system that is difficult to fabricate. References [1] and [2][1] suffer from all the problems associated with [3]–[9] and additionally raise a serious concern in that logic signal cannot propagate unidirectionally from the input to output since the input is linked to the output by Coulomb interaction which is bidirectional (reciprocal). Bidirectional flow of signal does not distinguish between input and output terminals [7]–[15]. There are additional problems in this paradigm such as the lack of isolation between input and output and the problem of unbalanced logic gates which tends to generate wrong answers for certain configurations of gates [16].

In this paper, we propose and demonstrate a novel paradigm for nanoelectronic implementation of Boolean logic that can potentially eliminate the above drawbacks. In this implementation, each computational cell contains nanoscale metallic dots formed into regimented, two-dimensional (2-D) arrays on an active substrate which possesses a nonlinear, nonmonotone current-voltage characteristic. Logic functionality is achieved through charge exchange between the metallic dots. In order to demonstrate the feasibility of making these circuits, we also outline a fabrication strategy which exploits chemical self-assembly/self-organization techniques to realize the nanoscale elements and the interelement connections. It appears that all of the individual components of the enabling technology have already been demonstrated and what remains to be achieved is the combination of these components to synthesize the circuits described in this paper. While this is certainly a nontrivial task, it seems that proof of concept demonstra-

Manuscript received November 5, 1996; revised January 14, 1997. This work was supported in part by the Army Research Office under Grants DAAL03-G-0144 and DAAH04-95-1-0586, and in part by the Defense Advanced Research Projects Agency under the ULTRA program Contract 35918-OH.

V. Roychowdhury is with the Electrical Engineering Department, University of California, Los Angeles, CA 90095 USA (e-mail: vwani@ee.ucla.edu).

D. B. Janes is with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA (e-mail: janes@yara.ecn.purdue.edu).

S. Bandyopadhyay is with the Department of Electrical Engineering, University of Nebraska, Lincoln, NE 68588-0511 USA (e-mail: bandy@engrs.unl.edu).

Publisher Item Identifier S 0018-9219(97)02729-1.

[1]This scheme lacks unidirectionality and does not work (see [7]–[15] for a discussion of this issue). Recent modifications of this scheme, whereby each cell is accessed and clocked individually to realize adiabatic switching, provides unidirectionality but results in exceedingly slow speed. There is still the lack of isolation between input and output. Furthermore, it is not clear how "logically irreversible" gates described in this paradigm can be switched adiabatically.

0018–9219/97$10.00 © 1997 IEEE

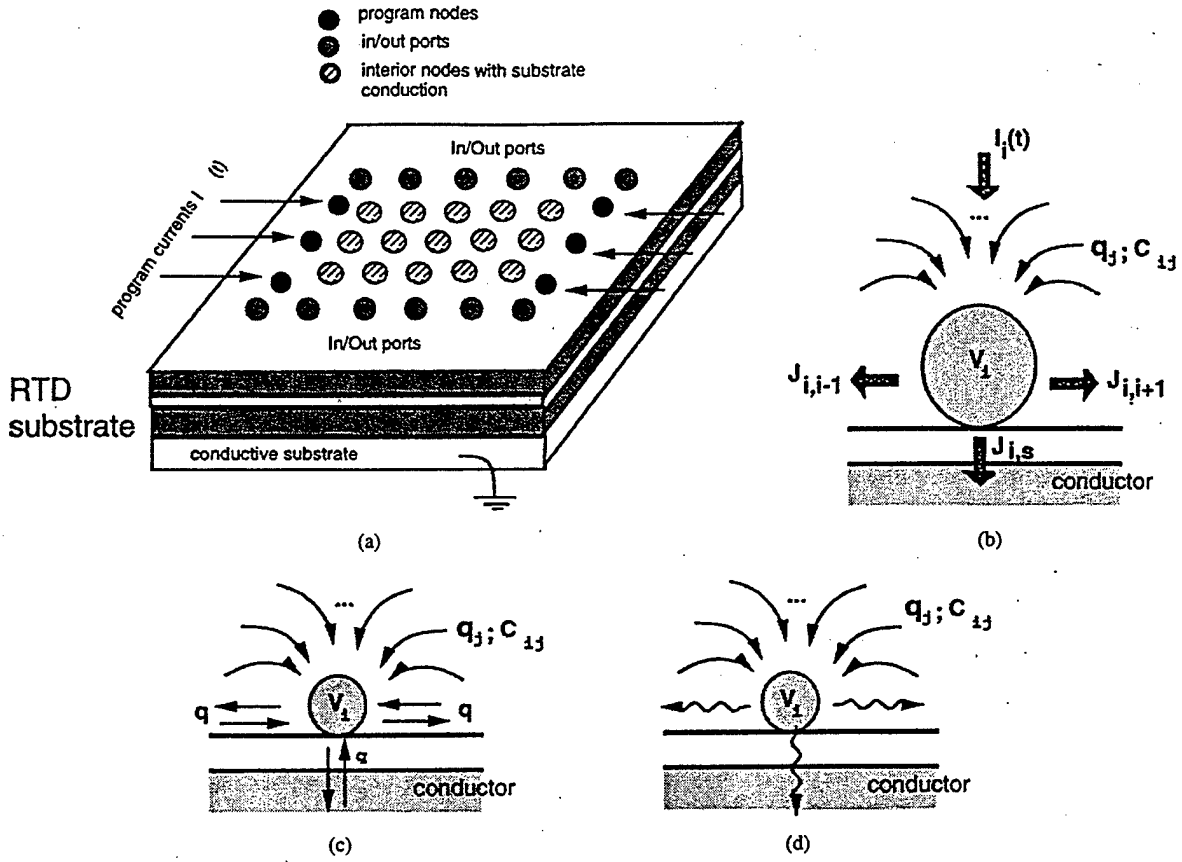



**Fig. 1.** (a) A generic array of metallic islands deposited on a nonohmic layer which has been grown on a conductive substrate. It is assumed that all islands have direct conductive links to nearest neighbors. A subset of the islands serve as program nodes which are driven by external current sources. Another subset of islands serve as input/output ports, and the remaining islands introduce complexity to the system through nonlinear conductive links to the substrate. The substrate nonlinearity could arise, for example, from resonant tunneling which causes the substrate current–voltage characteristic to have a *nonmonotonicity*. (b)–(d) Cross-sectional view showing island $i$, whose potential $v_i$ is influenced by the charges $q_j$ in the rest of the network through a capacitance matrix. The different transport mechanisms lead to a rich variety of computational capabilities.

tions of nanoscale computational cells using the proposed architecture will be possible in the relatively near future.

This paper is organized as follows. In Section II, we describe the paradigm, and present theoretical and simulation results to demonstrate the functioning of the logic gates. The rest of the sections are devoted to describing possible routes to fabrication.

## II. Architecture: Array of Quantum Dots on a Resonant Tunneling Diode

### *A. Basic Computing Block: Background and Motivation*

The basic computational block for our logic scheme can be realized by assembling a 2-D periodic array of nanometer-sized metallic islands with resistive nearest neighbor interconnections, on a substrate whose current-voltage characteristic has a *nonmonotonic* nonlinearity. The simplest choice for the substrate is a resonant tunneling diode (RTD). This system is shown in Fig. 1(a). All our circuits utilize this basic structure.

We have already shown that different types of transport nonlinearities in the interdot connections can lead to a wide variety of global activities in this system; moreover, it is possible to impart computational interpretations to these activities [14], [17]. Additionally, we have shown that within a classical circuit theoretic model (where electric charge is considered a continuous variable), nonmonotone nonlinearities of the substrate conduction can yield global associative memory effects and lead to image processing capabilities [10], [14], [17]. We have also shown that this interpretation remains valid even when granularity of charge (or single-electron effects) come into play, provided that the effective capacitance of the islands is not too small. Finally, we have investigated networks of islands in which the sole nonlinearity in charge transport arises from single-electronics (Coulomb blockade), rather than any inherent nonlinearity in the molecular wires or substrate conduction. These networks exhibit associative memory effects, as well as the ability to yield approximate solutions to certain NP-complete optimization problems such as the traveling salesman problem [10], [17].

Our past work has shown that the functionality of these networks is determined by the type and versatility of the interconnects. Increased functionality can be achieved at

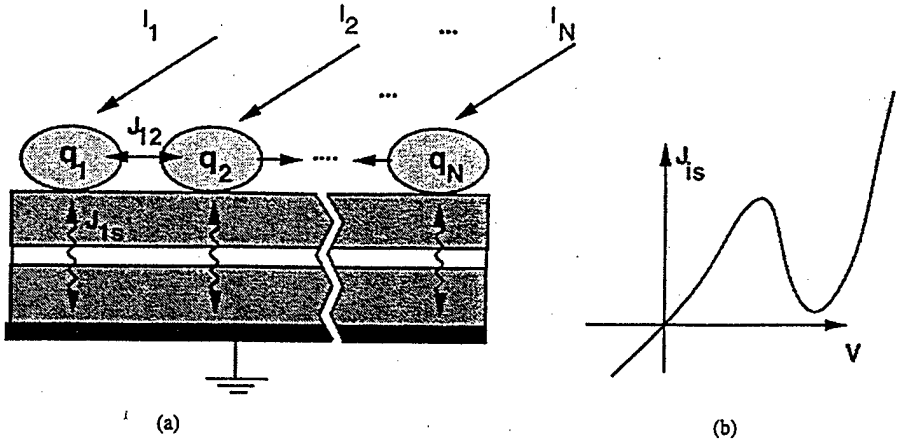



Fig. 2. (a) A 1-D array of islands, whose dynamics is described by (1). (b) A nonmonotone substrate nonlinearity of the kind shown here, is the minimal condition for the realization of collective effects.

the cost of a commensurate increase in the fabrication complexity involved in realizing the underlying networks. Hence, a natural question to ask is: what are the minimal requirements that the lateral interconnects must satisfy so that the underlying computational unit displays useful computational properties?

In this paper, we provide an answer to the above question, and show that the architecture shown in Fig. 1 can realize basic logic functions such as AND gates, OR gates, and multilevel circuits comprising these gates using the simplest of interconnects. We need just nearest-neighbor resistive links, and at certain prespecified locations that demarcate gate boundaries, we need rectifying or diode links.

### B. Bistability

In order to make logic gates and circuits, one first needs to demonstrate a bistable device, i.e., a device that is capable of assuming only two possible stable states, one of which can be interpreted as a binary 0 and the other as a binary 1. We shall show later that a bistable device by itself is not sufficient for realizing logic circuits (one also needs to impose "unidirectionality," a issue that is addressed later in this paper), but bistability is a necessary ingredient. In this section, we demonstrate how bistable devices can be generated using the computational prototype described in Fig. 1.

*1) Networks with Multiple Stable States:* First consider a simple one-dimensional (1-D) version of the network, as shown in Fig. 2. Here we designate the current between islands $i$ and $j$ as $J_{ij}$, and the current between island $i$ and the grounded substrate as $J_{is}$. We can then write from Kirchoff's current balance condition that,

$$\frac{d}{dt}q_i = C_i\frac{d}{dt}v_i = -\left[J_{is}(v_i) + \sum_{j\neq i} J_{ij}(v_i - v_j)\right] + I_i(t) \quad (1)$$

where $I_i(t)$ is the driving current, $q_i$ is the charge, $v_i$ is the potential, and $C_i$ is the capacitance, with the subscript $i$ indicating the relevant island.

Qualitatively, the multiple solutions to the nonlinear system of equations which results from (1) under steady-state conditions (i.e., $dv_i/dt = 0$ for all $i$) will be taken as a set of memory states (or equilibrium points) which can be programmed by properly choosing the current biases $I_i$, and the transport functions $J_{ij}$ and $J_{is}$. The current biases $I_i$ will be assumed to be either time independent, or slowly varying on time scales over which the network relaxes into its memory states. If the network is begun at time $t = 0$, with a certain initial condition $v_i = q_i/C_i$ for all $i$, arranged, for instance, by an initial impulse of charges $q_i$ at each of the islands, then the network will evolve toward the closest memory state, as programmed by the current biases $I_i$.

In [14], [17], we demonstrated two important properties of this network: 1) it is globally stable and 2) the nonmonotonic nonlinearity causes *multiple* (more than one) equilibrium points. The number of equilibrium points can be programmed by appropriate choice of the inter-island resistances. For binary logic, we need only *two* equilibrium points.

We next provide numerical results for multiple equilibrium points that can be observed in our networks by making different choices of $J_{ij}$'s and $J_{is}$'s. Our example, based on two dots, is illustrated in Fig. 3(a), in which the two islands are coupled with each other through a linear resistor $R_{12}$ and a capacitor $C$. In addition, we assume that each island is *individually* coupled to the substrate through a capacitance $C_0$, and a nonlinear resistor whose nonmonotonic current-voltage characteristic is shown in Fig. 3(b). This characteristic approximates that of a RTD. Since the results presented here are numerical in nature, we have minimized the number of parameters by choosing a piecewise linear function for the characteristic, in which each of the three

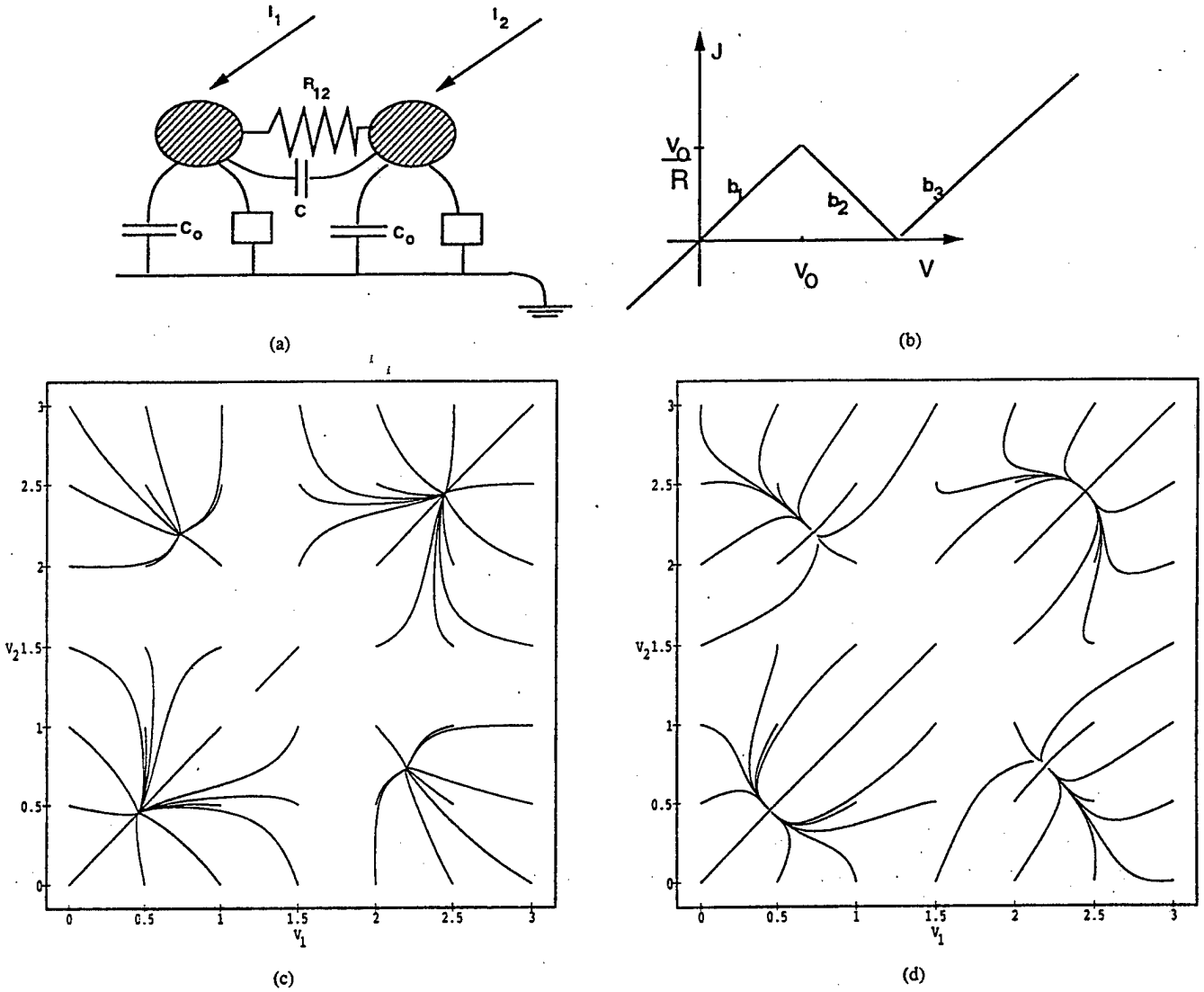
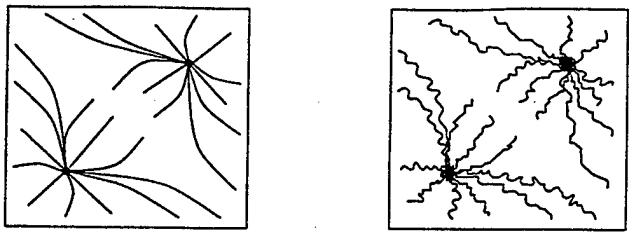



**Fig. 3.** (a) A network of two islands that are laterally coupled via a linear resistance $R_{12}$ and a capacitance $C$. Each island is also coupled vertically to the substrate via a capacitance $C_0$, and a nonlinear resistive element. (b) The substrate nonlinearity is modeled as a piecewise linear function. We can show theoretically that more realistic nonlinearities due to resonant tunneling, for example, will yield qualitatively the same dynamics as generated by this nonlinearity. The three segments of the function are denoted by $b_1$, $b_2$, and $b_3$. The other two figures show the phase portrait for two-island system in the continuous-charge model, in which the voltage axes have been scaled with respect to $v_0$. The parameters are $R_{12} = 5R$, and $I_0 = v_0/2R$. (c) With only substrate capacitance $C_0 = 1$, and (d) with identical mutual and substrate capacitances: $C = C_0$.

segments have slopes of the same magnitude $1/R$

$$J_s(v) = \begin{cases} \dfrac{v}{R} & v < v_0 \\ \dfrac{2v_0 - v}{R} & v_0 \le v \le 2v_0 \\ \dfrac{v - 2v_0}{R} & v > 2v_0. \end{cases} \tag{2}$$

The equilibrium points of this two-node system can be found by setting $\dot{v}_1 = \dot{v}_2 = 0$, and the stability properties of those equilibrium points can be ascertained by examining the eigenvalues of the system matrix near the equilibrium points [18], [19]. Provided that the driving current $I_0 < v_0/R$, each island potential can in principle be on any one of the three branches of the nonlinear function $J_s(v)$ shown in Fig. 3(b). We can, however, show that the system will be unstable if either island is operated on branch $b_2$ which has a negative differential resistance. So, it follows that this system can have at most four different globally stable points, since each island is restricted to being on either branch $b_1$ or $b_3$. Two of these stable points are trivial ones corresponding to both islands operating on the same branch: either $b_1$ or $b_3$. We have determined a necessary and sufficient condition for the existence of all four "memories," and the corresponding phase portraits are shown in Fig. 3. The figure shows that if the interdot resistance ($R_{12}$) is large enough, then all the four possible equilibrium points are stable. In that case, this system can be used as an associative memory.



(a) (b)

Fig. 4. (a) The phase portrait of the two-island system shown in Fig. 3, when the inter-island resistance is small, e.g., $R_{12} \leq R$. The whole system now behaves as a bistable device with only two stable states. (b) The phase portrait of the same system when single-electron effects and Coulomb blockade are taken into consideration.

*2) Networks with Two Stable States:* Now, if instead of an associative memory application with multiple stable states, one needs a device for binary logic functions, then we have to demonstrate the existence of only two stable points. We consider the same two-island example as described in (2) and Fig. 3. However, instead of choosing a large interdot resistance (i.e., $R_{12} = 5R$), we choose a small interdot resistance (e.g., $R_{12} = R$). Our analysis and simulations (see Fig. 4) show that with small inter-island resistance, instead of having four stable points, the two-node system will have exactly two stable states. These two stable states correspond to the voltages on both islands being on branch $b_1$ or branch $b_3$.

In general, one can consider any large network comprising a collection of $n$ islands [as depicted in Fig. 2 and characterized by (1)] and show that the whole network behaves as a single bistable device when the lateral interconnects have small enough resistances. A proof of this statement follows easily from the preceding analysis of the two-island case, and will be skipped here.

*3) Limits to Continuous Models—Single-Electron Effects:* We now study the same networks, with the assumption that single electron effects have become pertinent either due to the lowering of temperature, or due to the physical scaling of the metallic islands down to $d \sim 10$ nm. Single-electron effects will become relevant when the inter-island resistances exceed $h/e^2$ ($\approx 12$ k$\Omega$) and when the change in potential $\delta V = q/C$, associated with the addition of a single charge $q$ to an island, becomes comparable to $kT/q$, the thermal potential. The latter condition can be met even at room temperature if islands with an effective capacitance smaller than $C \approx 5 \times 10^{-18}$ F are fabricated. In fact, as described in the experimental section of this paper, we have already observed room temperature single-electron effects in the quantum-dot arrays that we have synthesized [20], [21].

Single electron dynamics are characterized by discrete tunnel events, and differential equation systems cannot be used to model the arrays. We have developed a Monte Carlo simulation technique [22], [23] for the simulation of a current biased network of islands shown in Fig. 2(a). This simulator has been described in [17]. The results obtained from this simulator show the following features [see also Fig. 4(b)]: For large enough values of the effective substrate capacitances ($C_0$), the effect of the single electron dynamics is marginal and the *system is adequately described by the continuous charge models*. Thus, even arrays of metallic islands with diameters of few nanometers can exhibit the same associative memory and bistability properties studied in the context of continuous dynamics.

### C. Networks for Logic Gates and Functions

The preceding discussions clearly show how a collection of resistively coupled quantum dots with an RTD substrate, can behave as a single bistable device. For the purposes of this section, we shall represent the computational structures described in Figs. 1(a) and 2(a), simply as an array of dots; it will be assumed that the quantum dots are laterally connected by sufficiently low-impedance links, and that there is an RTD substrate underneath the array. The results presented in this section are based on the outputs of the simulators that we have developed for solving the continuous-time (1), as well as, based on the Monte Carlo simulator designed to study the behavior of the networks under single-electron dynamics.

Given such an array, an implementation of a logical OR gate is demonstrated in Fig. 5. We will assume "positive logic" so that the high voltage state corresponds to logic level 1 and the low voltage to logic level 0. The diodes at the boundary provide isolation between the two inputs $A$ and $B$. The whole array is initialized to a 0 state, and if any of the inputs is a 1 then the array switches its state from a 0 state to 1. An analogous realization of an AND gate is shown in Fig. 6.

An example of the realization of a two-level OR/AND circuit is shown Fig. 7. The figure shows the results of our numerical simulations and establishes how the individual gates can be integrated into realizing Boolean circuits. The first level of the circuit shown in Fig. 7 consists of two 2-input OR gates. In the first step, the OR gates are allowed to operate while the AND gate in the second stage is disabled. Next, the AND gate is enabled by a *clock* pulse and the two OR gates drive the AND gate. Multilayered logic circuits can be also realized by propagating the signals stage by stage. That is, successive levels or stages of the circuit are activate sequentially allowing the signal to propagate from the input to the output of the circuit.

Note that we have not provided a realization of an inverter (or a NOT gate) in our scheme. That is because, without loss of generality, we can assume that for every binary input $A$, its inverted value $\overline{A}$ is also available as an input. This is the usual assumption made in many currently available schemes for logic implementations, such as the programmable logic arrays (PLA's), where all the variables and their complements are available as inputs. The inversion of the inputs can be also done at the boundaries of our computational block using nanoscale single electron transistors (SET's). It can be easily shown from the basic principles of Boolean logic that if input variables and their complements are available, then a two-level OR/AND circuit is universal.

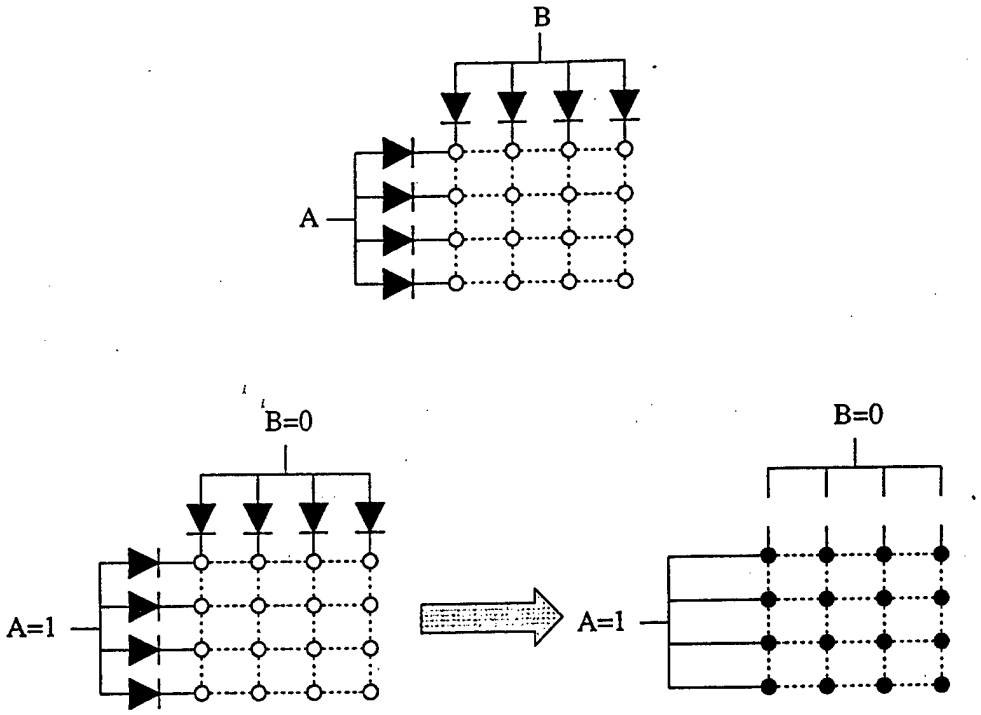



**Fig. 5.** A schematic description of the realization of a two-input OR gate using arrays of metallic dots deposited on a nonohmic layer. The metallic dots are interconnected with each other by resistive links, and to the inputs by *rectifier* (or *diode*) links. The computation starts by initializing the individual dots to a *low* state, and if any of the inputs is a logical 1, then all the dots will switch to a *high* state (represented by •).

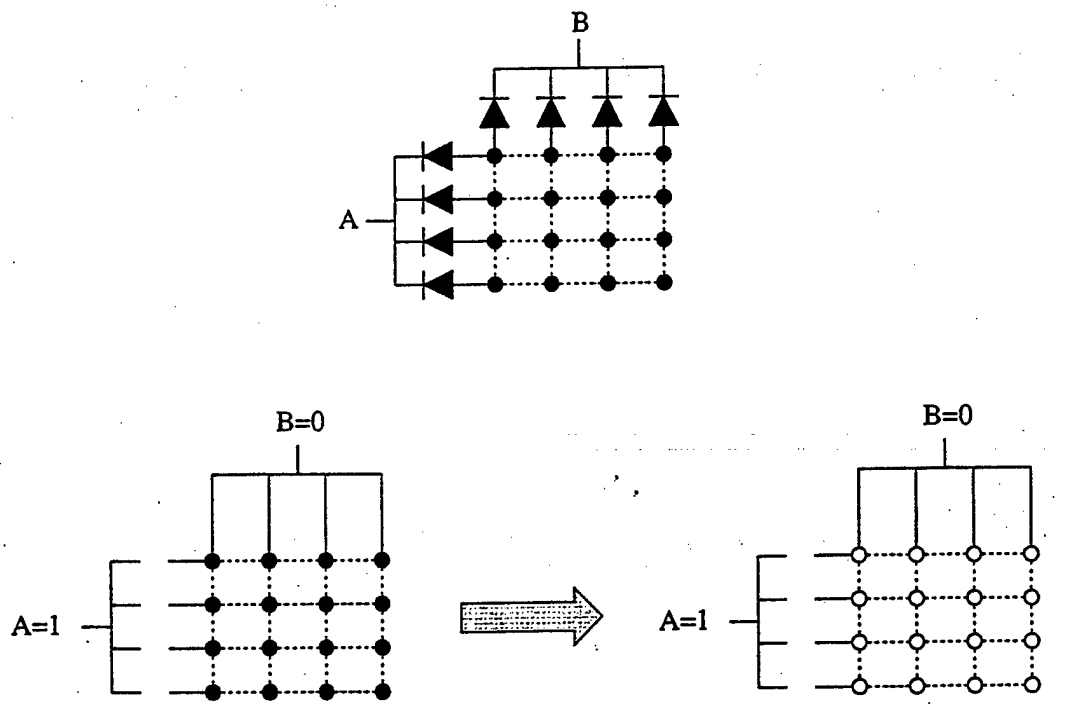



**Fig. 6.** A schematic description of the realization of a two-input AND gate using arrays of metallic dots deposited on a nonohmic layer. The metallic dots are interconnected with each other by resistive links, and to the inputs by *rectifier* (or *diode*) links. The computation starts by initializing the individual dots to a *high* state, and if any of the inputs is a logical 0, then all the dots will switch to a *low* state.

### *D. Basic Architectural Issues*

In this section, we highlight several architectural issues involved in the scheme that we have developed.

*1) Unidirectionality and Pipelining:* As described in the introduction, a major drawback of the proposals for logic implementations based on bistable devices without direc-

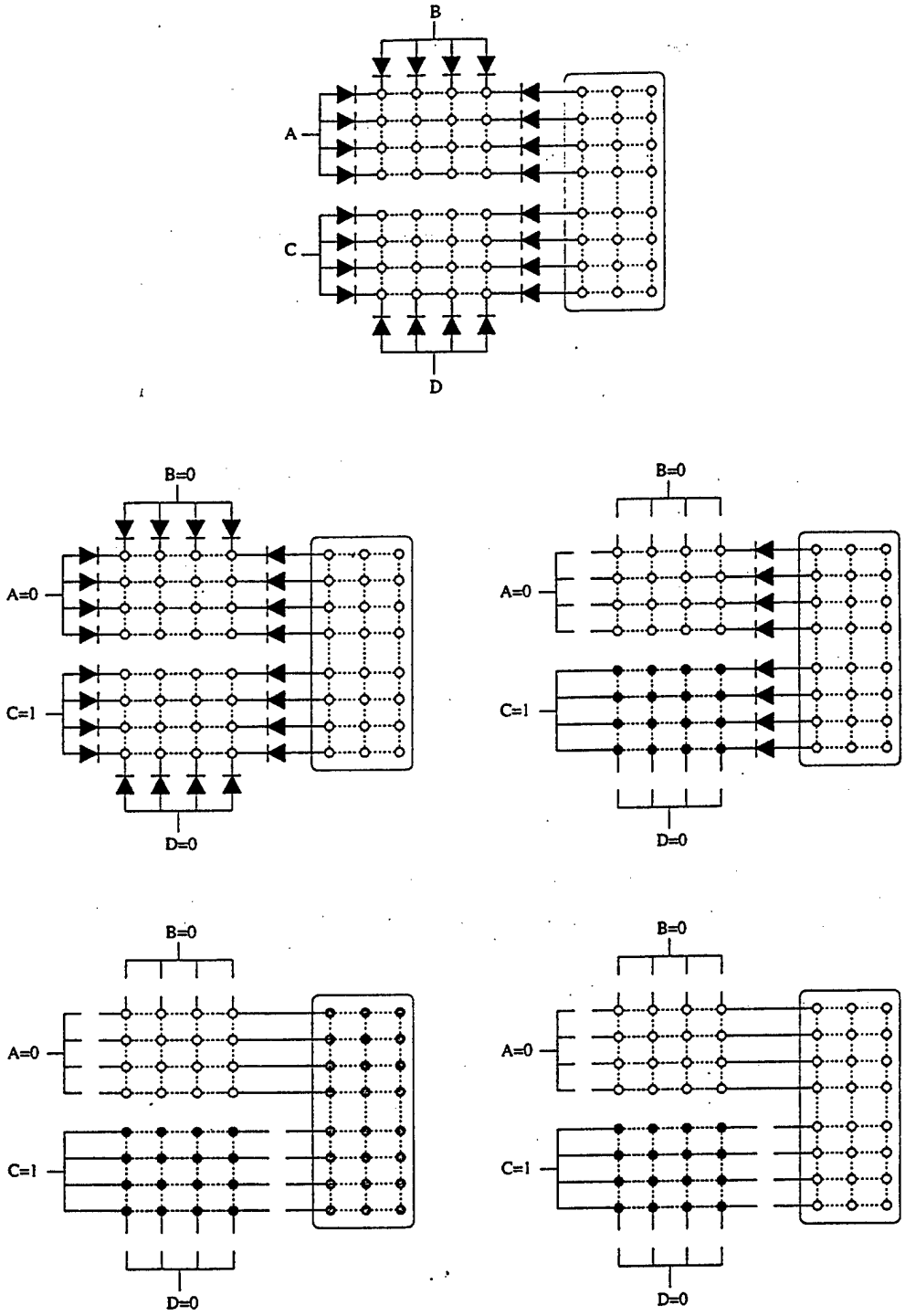



**Fig. 7.** A schematic description of the realization of a two-level OR/AND circuit using arrays of metallic clusters deposited on a nonohmic layer. *Rectifier* (or *diode*) links delineate gate boundaries. The computation starts by initializing all the individual dots to a *low* state. This allows the OR gates to complete their respective computations. This step is followed by an initialization of the AND part of the array to a *high* state. The final state of the AND part will be determined by the already computed states of the OR gates.

tionality [1] lies in their failure to ensure propagation of the logic signal from the input to the output. That is, if two bistable devices are connected together in series, then the one closer to the input terminal should drive the other, and not the other way around. This will ensure that the signals propagate from the input to the outputs and the whole system does not get stuck in metastable states [12], [13], [15].

As illustrated in Fig. 7, in our scheme, unidirectional signal propagation from the input to the output is effected through a *clocking mechanism*. For example, when the OR gates in the first layer are operational, the AND gate

in the succeeding stage is disabled. Once the OR gates have reached stable states, then a clock pulse is applied to initialize the AND gate. This scheme avoids the potential problem of the AND gate acting as an input to the OR gates, and instead allows the OR gates to drive the AND gate. The same strategy can be extended to the case of multi-level circuits. Such a scheme of multiphase clocking ("push-clock" and "drop-clock") is also used in conventional charge-coupled devices (CCD's) [24].

Moreover, the multiphase clocking scheme will enable *pipelining* in our computational blocks. That is, every other level in a multilevel circuit can operate simultaneously, and a new set of inputs can be fed to the circuit every other clock cycle. This allows the implementation of a high throughput logic block.

*2) Fault Tolerance:* Any computational architecture at the nanoscale level should display inherent fault-tolerant properties [12], [13]. Nanostructure devices will probably have more variability in their characteristics than their earlier generation microstructure counterparts. Any scheme that ignores this fact and relies on every quantum dot being perfect will almost inevitably be impractical.

In order to ensure fault-tolerance in our scheme, we allow a cluster of islands to represent a gate rather than a single or just a few dots. Note that the size of the arrays for each gate (as shown in Figs. 5 and 6) can be varied depending on the state of the technology. By providing larger arrays one can increase fault tolerance. We have done extensive simulations of the two-level OR/AND circuit shown in Fig. 7 under different fault conditions. For example, our simulations show that the operation of the circuit remains unimpaired even if up to 15% of the dots fail. Similar behavior holds even if the sizes of the dots are varied up to 100%.

*3) Power Dissipation and Density Issues:* Given the high circuit density of the proposed computational architecture, it is important to consider the power dissipation of the cells. While dynamic power simulations are not available at this time, the static power dissipation can be estimated from the stable current and voltage states. For the case with optimally chosen current bias to each node, the static current level for both stable states of each island will be approximately the valley current level of the RTD mesas; the corresponding voltage levels will be approximately the valley voltage and zero voltage for the high and low voltage states, respectively. For a cell containing 5 nm diameter islands, the static power dissipated in the high voltage state is approximately 0.1 nW per island, while the static power in the low voltage state is approximately zero. This calculation assumes a peak current density of $10^4$ A/cm$^2$, a peak to valley current ratio of 10:1 and a valley voltage of 0.5 V, all consistent with RTD performance reported in the literature. A computational cell containing 100 islands would therefore dissipate 10 nW and would occupy about $1 \times 10^{-10}$ cm$^2$. To a first approximation, only half of the computational cells will be biased at a given time and only half of the nanoscale islands in the active cells will be in the high voltage state. Therefore, the static power dissipation is estimated to be 25 W/cm$^2$ for a computational circuit with $1 \times 10^{10}$ cells per cm$^2$.

For comparison, projected values for silicon CMOS circuits in the year 2007 are a dissipation of 600 nW per computational cell (logic gate) with a cell density of $5 \times 10^7$ cells/cm$^2$, corresponding to 30 W/cm$^2$ [25]. The CMOS power is primarily dynamic power, so detailed comparisons to the static predictions of the proposed architecture are probably not appropriate. However, the power estimations indicate that the proposed architecture has the potential for significantly reduced power dissipation per cell and comparable power densities, along with the potential for realizing higher functionality per cell and much higher cell densities.

We can also estimate an upper limit for the dynamic energy dissipation during a logic signal swing. This energy is $\sim CV^2$ where $C$ is the capacitance of an island and $V$ is the power supply voltage. Assuming $C = 1$ aF and $V = 1$ V, the energy dissipation (or power-delay product) is $10^{-18}$ J.

The switching delay is $\sim R_{link}(C_{\text{sub}} + C_{\text{island}} + C_{\text{link}})$ where $R_{\text{link}}$ is the resistance of the interdot resistive link, $C_{\text{sub}}$ is the substrate capacitance associated with the RTD's, $C_{\text{link}}$ is the capacitance between nearest neighbor dots, and $C_{\text{island}}$ is the capacitance of the metallic island. Assuming that the value of $R_{\text{link}}$ is 1 MΩ, a value consistent with the interdot resistances of metal dot arrays described in the next section [21], and all capacitances are on the order of 1 aF, we obtain a switching delay of 3 ps. Therefore, the dynamic power dissipated is about 300 nW per island. However, one should remember that this is an upper limit. Heat-sinking of 1000 W/cm$^2$ from a silicon chip was demonstrated more than 15 years ago [26].

## III. Nanoscale Fabrication

There are several key components required to fabricate computational cells of the proposed architecture, including definition and interconnection of the cell core, i.e., the area covered by a uniform array of metallic nodes, the rectifying interconnects which connect adjacent cell cores and the bias/clocking circuitry. This section provides a brief overview of synthesis techniques that could be employed to realize the required components and interconnections. The basic cell configuration will be briefly reviewed and the requirements for fabricating a logic element using this topology will be discussed.

It should be noted that one of the attractive features of the proposed architecture is its compatibility with fabrication techniques of the type described in this section. In this case, local interconnects between the nanometer scale elements can be uniform resistive links, i.e., it is not necessary to provide arbitrary interconnection paths at the level of the smallest circuit nodes. This feature can be realized with self-organized or self-assembled networks, which can be formed with a relatively high throughput and relatively low cost. The internal connections to each nanometer scale element in the computational cores consist of a bias current

line and a load with a nonmonotone, nonlinear current-voltage relationship. These elements provide a voltage bistability for each element, as well as the power required to switch. The specific functionality of the circuit comes from the intercell interconnections. Since these interconnections can be at larger scales than the nanometer scale nodes, it should be feasible to use lithographic techniques for their delineation and this can provide more arbitrary interconnections.

Since the architecture is intended to be scalable to nanometer dimensions, the underlying assumption of this discussion is that the individual metallic nodes will have dimensions in the range of 5–20 nm and that the computational cells will have dimensions (lengths and widths) in the range of 20–200 nm. The minimum feature sizes are well below those employed in conventional silicon circuits. While the small areas of computational cells would result in very large circuit densities (potentially approaching 1 T gates/cm$^2$), realization of circuits at this scale will require a significant shift in fabrication techniques from those used in conventional silicon processing. Conventional lithographic techniques, including capabilities such as electron beam lithography, might be capable of patterning at the scale of the (larger) cell regions but would not be practical for defining the individual metallic nodes which are much smaller. In contrast, chemical self-assembly or self-organizing techniques, i.e., techniques which exploit chemical affinities rather than lithographic techniques to arrange clusters or molecules, have been developed which can provide highly organized structures with minimum sizes and spacings on the order of 2–10 nm. While these techniques can provide uniform arrays of molecules or metallic islands, they do not appear to be capable of spontaneously forming specific nonuniform patterns or arbitrary interconnect configurations. The fabrication techniques best suited for realizing the structures required for the proposed computational cells will likely be a combination of chemical self-assembly techniques, to realize the networks of nanoscale metallic nodes within a computational cell, and lithographic techniques to define the cells and interconnect structures.

An equivalent circuit representing a 1-D slice through two computational cells is shown in Fig. 8. Each computational cell contains a number of nanoscale metallic nodes, as illustrated in the figure. Adjacent nodes within the cell are coupled resistively via intracell resistances ($R_1$), as well as capacitively via $C_1$. Each nanoscale node is connected to an active element, labeled "RTD" in the figure, which provide the required nonmonotone nonlinearity. The intercell diode connections are configured to realize specific logic functions. Each cell core is biased through a clock electrode; two or more clock phases are required to provide unidirectional propagation of logic signals. When enabled, the clock lines provide the bias voltage to the nanoscale nodes within a given cell. The bias resistance, $R_{bias}$ can be realized through a thin-film resistive layer.

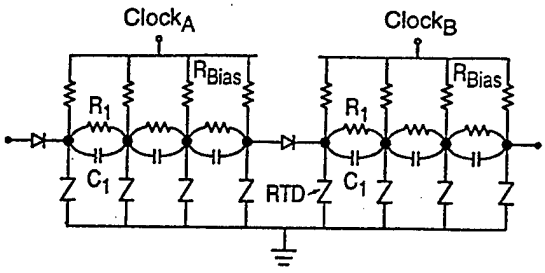



Fig. 8. Circuit schematic corresponding to a 1-D slice through adjacent computational cells. Indicated elements include nodes representing nanoscale metallic islands, bias resistors, bias/clock electrodes, elements with nonlinear nonmonotone current-voltage (provided by semiconductor mesas and labeled as RTD), and intercell rectifying links.

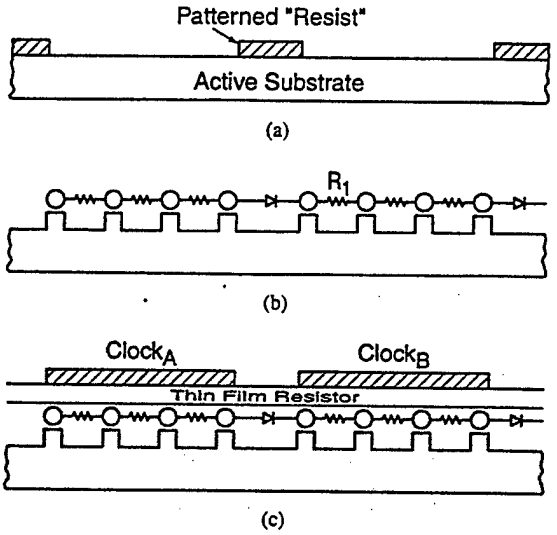



Fig. 9. Cross-sectional view of major steps in fabrication sequence for computational cells. (a) Lithographic patterning of regions in which computational cell cores will be deposited. (b) Self-organized deposition of arrays of nanometer scale metallic islands in patterned core regions. Internode resistive linking and patterning of semiconductor mesas, realized by chemical self-assembly and etching using metallic islands as "natural" masks, respectively, are also shown. (c) Deposition of thin film resistor layer for bias resistance and bias/clock electrodes for each cell, in regions defined by lithographic patterning.

A general synthesis approach can be formulated, based on self-organization and lithography techniques described later. The major steps in the synthesis procedure are illustrated in Fig. 9. The active substrate for this structure is a semiconductor heterostructure which contains layers designed to provide the required nonmonotone nonlinearity. For example, a double barrier RTD structure grown using molecular beam epitaxy (MBE) or metal-organic chemical vapor deposition (MOCVD) could provide an appropriate current-voltage relationship. The synthesis procedure begins with the lithographic definition (at a scale of approximately ten times the nanoscale node dimension) of regions which will contain the computational cell core regions, i.e., the regions in which arrays of nanometer diameter metallic islands are deposited. The arrays of metallic islands, along with the internode linking

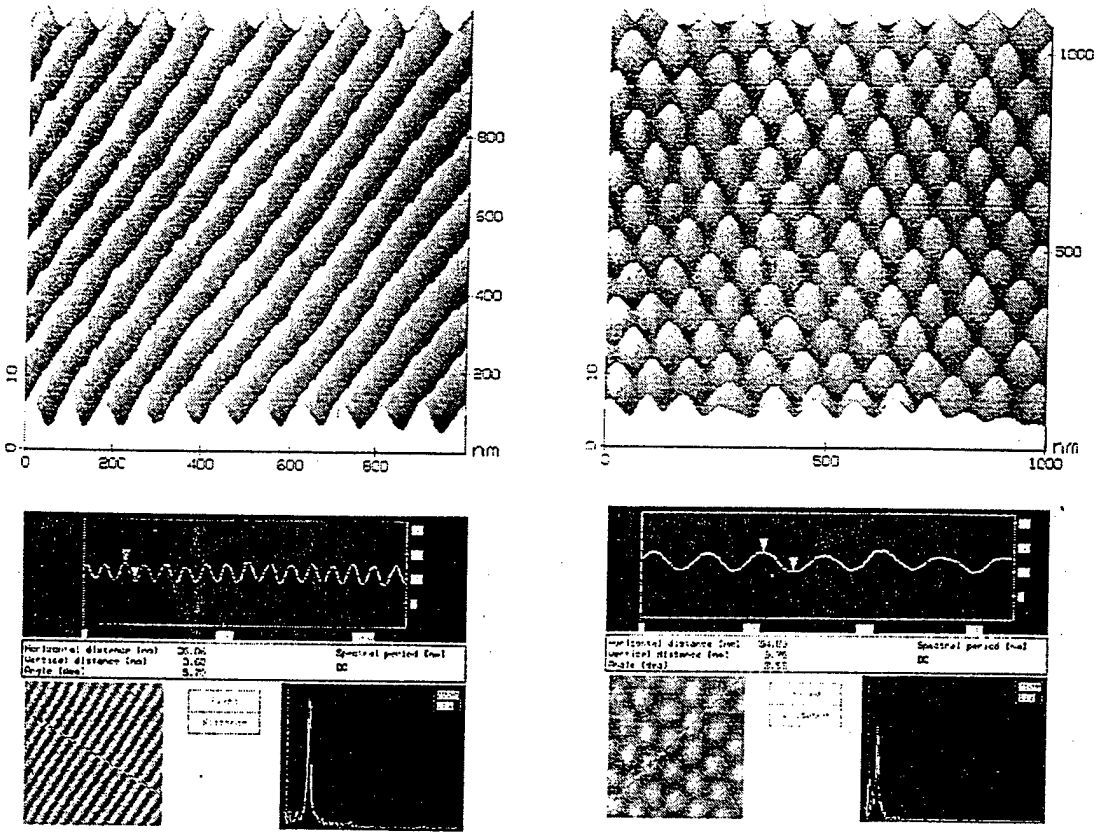



**Fig. 10.** Two typical patterns produced by electropolishing a thin film of aluminum. The first can be used to produce narrow metallic wire arrays and the last can be used to produce a hexagonal close-packed array of metallic islands. These figures are raw atomic force micrographs. Depth profile along arbitrarily selected directions on the surface are shown with well resolved peaks in the corresponding Fourier spectra indicative of the high degree of periodicity and ordering.

resistors, are deposited in these patterned regions using, for example, self-organizing techniques. The active substrate must be patterned to provide individual mesas for each of the metallic islands. In order to provide self-alignment of the semiconductor mesas with the nanoscale metallic islands, it is desirable to use the metallic island arrays as "natural" masks for etching of the mesas [27], [28]. Next, it is necessary to define the intercell coupling elements, namely the diodes. Finally, a thin film resistive layer, to provide effective bias resistors for each nanoscale metallic island, and bias/clock electrodes must be deposited over each computational cell. The deposition of electrodes for the polyphase clock requires another lithographic step, with registration to the cell definition lithography. Details of these procedures are described below.

### *A. Nanoscale Elements in Cell Core*

The primary features of the computational cell core are 1) a uniform array of nanometer scale metallic islands with controlled resistive coupling between islands, 2) low-resistance coupling of the metallic islands to semiconductor mesas with appropriate nonmonotone nonlinearity, and 3) a thin film resistive layer to provide bias connections to the metallic islands. Fabrication of the metallic nodes and active substrate connections within a cell have been described previously [14]. There are several techniques based on chemical self-assembly which can provide uniform 2-D arrays of metal islands suitable for cell core applications. Highly ordered, self-organized arrays of quantum dots and wires have been demonstrated using electropolishing of an aluminum film in an acidic solution of perchloric acid, butyl cellusolve, and ethanol [10], [27], [29], [30]. Fig. 10 shows examples of arrays of wires or dots formed by the electropolishing technique. The feature sizes of these patterns can be controlled over the range of approximately 20–100 nm by varying the temperature and composition of the electrolyte and possibly the voltage of electropolishing. The required nanoscale islands can be realized either by directly employing the patterned aluminum film or by pattern transfer to another material [14]. Resistive links between the islands can be provided by depositing a material such as low-temperature MBE grown GaAs on the surface to contact the islands.

Another independent approach that has been demonstrated involves fabricating large area, 2-D close packed arrays of 4 nm diameter gold clusters with controlled intercluster resistances [20]. In this technique, the arrays are formed from colloidal suspensions of neutral, encapsulated gold clusters. The intercluster resistance can be

controlled by exposing the array to conjugated organic molecules with end groups which bind to gold [31], [32]. An example of such a molecule is biphenyl-dithiol. The conjugated molecules bridge the gaps between clusters and result in a strong mechanical link between clusters and a controlled intercluster resistance [20], [21]. The measured in-plane conductance of linked cluster arrays formed using this procedure exhibit significant single electron charging effects at room temperature, consistent with the predicted capacitances for 4 nm diameter clusters and the predicted resistances for the conjugated organic linking molecules [20], [33]. The resistances of the organic molecules deposited between gold surfaces inferred from the array experiments and from scanning tunneling microscope experiments on individual clusters are in the range of 10–40 MΩ per molecule [20], [34]. Theoretical predications indicate that the resistance per molecule can be varied significantly by employing different conjugated organic molecules [33]; experimental evidence supports this prediction. While experiments to date have employed gold clusters, nanometer scale clusters of other metals have been synthesized using similar techniques and it is expected that 2-D arrays could be formed with other materials. In addition, alternate synthesis techniques yield clusters of semiconductor materials such as CdS and CdSe [35].

In order to provide effective coupling to the active substrate mesas, low resistance ohmic contact is required between the nanoscale metallic islands and the semiconductor mesas. The interface between the metallic nodes and the semiconductor surface layer must provide a stable, low-resistance contact without the need for alloying or other high temperature processing, since the self-assembly techniques used to form the metallic island arrays tend to be room temperature processes and nonreactive with the surface. In addition, post-deposition annealing of nanometer diameter islands could result in unacceptable size variation and shape distortion. A nonalloyed ohmic contact structure utilizing low-temperature grown GaAs has been demonstrated which should provide a suitable coupling layer for the node to active substrate interface [36]. This structure provides low-resistance ohmic contacts without annealing and also provides a surface which can remain free of significant oxidation even after hours of air exposure [36], [37]. In addition, this structure results in a controlled effective surface potential, without a significant surface depletion layer, as evidenced by the observation of a midgap band of states in a scanning tunneling spectroscopy experiment performed following air exposure [38]. Layers as thin as 2 nm are effective in providing this surface passivation effect and also allow thin $n$-type layers with very high activated donor densities to be realized immediately below this surface layer [36]. A relatively shallow etch can be used to selectively remove this low-temperature grown GaAs passivation layer, resulting in the depletion of the exposed surface and therefore providing effective patterning. The resulting surface is nearly planar, in contrast to high aspect ratio pillars often reported for small area device mesas. This near planarity is essential for subsequent fabrication steps, which may include self-assembly deposition of conjugated organic molecules for inter-island linking and the deposition of the thin-film resistive bias layers.

In the proposed cell cores, each nanometer scale metallic island needs to be connected to an individual semiconductor mesa. The requirements for registration between the islands and the mesas dictate a self-aligned technique. One possibility would be to use the self-organized metallic islands themselves as the masking elements, in a "natural" patterning technique comparable to that recently reported for 20–50 nm islands of gold evaporated onto a semiconductor substrate [28]. Other possibilities include using a masking layer, formed using a lithographic technique or a self-organized array, to define the semiconductor mesas and then to nucleate a metallic island on each mesa. Mesa isolation is achieved either by reactive ion etching, or by using the metal islands as a mask and photo-oxidizing the exposed areas in ultraviolet (UV) light.

### *B. Cell and Interconnect Patterning*

While self-assembly techniques can realize highly uniform arrays of metallic islands over relatively large area, the delineation of cell cores, intercell connections and bias/clock lines will require patterning techniques with minimum feature sizes approximately equal to the cell core dimension. For a computational cell in which the nanoscale metallic islands are 10 nm in diameter, the required minimum feature size would be 20–100 nm. High resolution lithography techniques such as electron beam or X-ray lithography can realize features in this range, but these techniques can induce surface damage and generally require photoresist materials that may be incompatible with the self-organizing techniques which may be used for definition of the nanoscale elements. There are, however, several demonstrations of patterning techniques which can provide the required feature sizes and which are compatible with the definition of self-assembled nanoscale arrays in patterned regions. Self-assembled monolayers (SAM's) of organic molecules, the most famous example of which is alkanethiol on gold [39], [40] could potentially be used to define areas for deposition or etching. SAM's on various surfaces, including semiconductors such as Si and GaAs, have been used to passivate the surface or to provide controlled surface characteristics such as hydrophobic/hydrophilic regions [41]. SAM's have also been employed as electron beam or scanning tunneling microscope resists [42]. In addition, deposition of SAM's using elastomer stamp pads or patterning of photoresist materials using nanoimprint lithography techniques can potentially provide low cost, high throughput patterning, provided that issues such as registration are addressed [41], [43]. The characteristics achievable with various SAM's, including small thickness (typically 1–2 nm), uniformity and chemical affinities, makes them well suited for applications involving nanoscale islands. Deposition of gold nanoclusters on organic layers patterned by electron beam lithography has been reported,

indicating that deposition within patterned regions is feasible [44]. In contrast, conventional photoresist materials may prove unsuitable as patterning templates for nanoscale arrays due to the residual surface contamination left after development and the relatively large thickness of the resists, in comparison with the dimensions of the nanoscale metallic islands.

While the fabrication of the proposed computational cells will require several significant development efforts, proof of concept experiments exist for most of the major steps outlined above. One area of difficulty is the need for intercell rectifying elements with relatively low turn-on voltages. There has been a great deal of interest in molecular rectifiers since the paradigm of a donor-$\pi$ system, $\sigma$ bonded tunneling bridge and an acceptor-$\pi$ system was proposed [45]. Theoretically speaking, rectification should occur in nearly any molecule with sufficient asymmetry in the one-electron energy levels [46]. Molecular rectification has been demonstrated unambiguously in Langmuir–Blodgett films [47] although, to our knowledge, there has been no report of rectification by a single molecule as opposed to Langmuir–Blodgett films. Reports of rectifying characteristics for molecules using a TTF/TCNQ donor/acceptor configuration deposited between metal plates [48] indicate that rectification is possible in elements which are compatible with the dimensions and chemical characteristics of the proposed synthesis techniques. However, the reported current densities ($\mu$A/cm$^2$) have been well below the levels required for operation in this application [48]. Assuming that a suitable molecular rectifier is available, it will be a different molecule than the one needed to provide the resistive links between the metallic islands within a cell core. Two levels of molecular self-assembly in patterned regions will therefore be required for the two interconnect functions, namely intracell resistive connections and intercell rectifying connections Other alternatives, including rectifiers incorporated in the semiconductor active layers, can also be explored.

### C. Alternate Molecular Diodes

The molecular diodes described in the previous section are incapable of handling sufficiently large current densities. To overcome this drawback, we can explore two alternate routes for synthesizing diodes to link selected quantum dots. These techniques result in vertical diodes between a molecule and an underlying semiconductor substrate. Consequently, they are not immediately suitable for synthesizing lateral rectifying connections between metallic islands. However, these diodes can carry significantly higher current densities, well in excess of $10^6$ A/cm$^2$ [49] and are therefore attractive. Needless to say, significant further research needs to be done in this area before these vertical diodes can be modified to provide lateral connections. We describe these diodes below.

*1) Scanning Tunneling Tip Induced Molecular Decomposition:* The combination of a CVD source molecule and a scanning tunneling microscope (STM) provides a way to "direct write" nanometer-sized diodes on a suitable substrate. In this process, the energy-tunable electron flux from the highly positionable STM-tip is used to decompose the appropriate regions of a deposited layer of CVD source molecules on the substrate. In these regions, a heterojunction diode is formed between the decomposition product and the underlying substrate. This is followed by annealing to remove any unwanted species or undecomposed molecules through thermal desorption [50]. The final product is a designed array of local regions with altered composition forming heterojunction diodes. Organometallics and main group carboranes are potential candidates for CVD source compounds to be used in this approach [51]–[55]. For this process to be truly a selective area process, the source compound must adsorb molecularly rather than dissociatively [50]. Carpinelli *et al.* [50] have shown that such a molecular precursor state exists over a limited coverage range at room temperature for the icosahedral cage molecule *closo*-1, 2-dicarbadodecaborane ($C_2B_{10}H_{12}$) on Si. They studied the surface reaction of Si(111)-(7 $\times$ 7) reconstruction with the above molecular species using an STM. The current-voltage characteristic of the $B_5C$/Si(111) interface (measured with STM) showed anisotropic rectifying behavior indicating that this is truly a process for "direct-writing" nanometer-sized heterojunction diodes.

The above technique is a "direct-write" process and hence serial in nature. Each wafer has to be patterned one at a time resulting in slow throughput. A faster technique is to make the process compatible with masking and exposure. This is realized as follows. Decomposition of *closo*-1, 2-dicarbadodecaborane can be achieved by exposing the molecule to either UV or X-ray radiation. UV decomposition of organometallics or cluster source molecules offers better selective chemistry than X-ray, partly because decomposition by X-ray is due to secondary electrons [56], [57]. However, X-ray is superior in that smaller feature sizes can be delineated since the wavelength is shorter. There have been definite indications that a $B_5C$/Si(111) heterojunction diode can be fabricated by synchrotron (X-ray) induced decomposition of *closo*-1, 2-dicarbadodecaborane on Si [58]. This can lend to a projection lithography technique for synthesising $B_5C$/Si(111) heterojunction diodes. Since this process involves exposure of the chemisorbed species to the X-ray source through a mask, it will result in a high throughput.

## IV. Conclusion

We have described a novel architecture for realizing nanoelectronic logic gates and circuits based on arrays of metallic dots assembled on a double-barrier RTD substrate. Resistive connections link nearest neighbor dots within a logic cell and rectifying connections link the cells. The use of minimal complexity in the nanoscale interconnection elements should reduce the processing demands for this architecture. We have presented a theoretical analysis of this architecture, and provided results of simulations to exemplify logic operations. We have shown that the cells

can exhibit bistable voltage states in both the continuous charge and discrete charge (single electron) regimes. Topologies for realizing OR and AND gates as well as somewhat more complex Boolean circuits have been described. With appropriate cell configurations and clocking sequences, unidirectional propagation, fault tolerance, and manageable power dissipation levels can be achieved in the logic circuits. These systems, if realized, could result in extremely dense circuits with high operating speeds and sufficiently low power dissipation for ultra large scale integration.

A fabrication procedure for cells and interconnections of the architecture has been described. The fabrication incorporates self-organization and "natural" patterning techniques to realize the nanometer scale elements, internode resistance, and self-aligned semiconductor mesas required in the cell cores. Cell areas, bias/clock electrodes and interconnect areas are defined using conventional nanolithographic techniques which are compatible with the chemical self-assembly based techniques used for the nanometer scale elements. Although proof of concept experiments exist for the major technologies involved in the fabrication procedure, integration of the overall fabrication sequence, and control of various element values undoubtedly present significant challenges. Nonetheless, these challenges may be overcome with time. Moreover, it is anticipated that this technology could be readily adapted to fabrication of other computational circuits with nanometer scale elements. Thus, we conclude that these architectures and alternate fabrication methodologies for synthesizing them deserve serious attention.

## ACKNOWLEDGMENT

The authors would like to acknowledge the pioneering work on synthesis, characterization, and modeling of linked cluster networks by our collaborators at Purdue University, including Prof. S. Datta, Prof. R. P. Andres, Prof. R. Reifenberger, and Prof. C. P. Kubiak, as well as J. D. Bielefeld, J. I. Henderson, S. Hong, and V. R. Kolagunta. The proposed synthesis techniques would not be possible without the ideas and efforts of these individuals, along with Purdue collaborators D. McInturff, T.-B. Ng, and Prof. J. M. Woodall and Prof. M. R. Melloch. The authors are also indebted to co-investigators A. E. Miller and H.-C. Chang at the University of Notre Dame for collaborations leading to the pioneering work on the electrochemical synthesis of highly self-ordered patterns on aluminum (Fig. 10). This work also drew on the expertise of G. Banerjee, D.-F. Yue, and V. Yuzhakov. The authors are indebted to Prof. M. A. Ratner for insightful discussion on molecular rectifiers and to Dr. P. Balasingham for discussions about nanoelectronic networks.

## REFERENCES

[1] P. Bakshi, D. Broido, and K. Kempa, "Spontaneous polarization of electrons in quantum dashes," *J. Appl. Phys.*, vol. 70, p. 5150, 1991, and P. Bakshi (unpublished).

[2] C. S. Lent, P. D. Tougaw, W. Porod, and G. H. Bernstein, "Quantum cellular automata," *Nanotechnology*, vol. 4, p. 49, 1993.

[3] K. K. Likharev and A. N. Korotkov, "Single electron parametron: Reversible computation in a discrete state system," preprint.

[4] A. N. Korotkov, "Wireless single-electron logic biased by alternating electric field," *Appl. Phys. Lett.*, vol. 67, p. 2412, 1995.

[5] R. A. Kiehl and T. Ohshima, "Bistable locking of single electron tunneling elements for digital circuitry," *Appl. Phys. Lett.*, vol. 67, pp. 2494–2496, 1995.

[6] T. Ohshima and R. A. Kiehl, "Operation of bistable phase locked single electron tunneling logic elements," *J. Appl. Phys.*, vol. 80, pp. 912–923, 1996.

[7] S. Bandyopadhyay, B. Das, and A. E. Miller, "Supercomputing with spin polarized single electrons in a quantum coupled architecture," *Nanotech.*, vol. 5, pp. 113–132, 1994.

[8] S. S. Molotkov and S. N. Nazin, "Single electron spin logical gates," *JETP Lett.*, vol. 62, p. 272, 1995.

[9] ——, "Single electron computing: Quantum dot logic gates," *Zh. Eksp. Teor. Fiz.*, vol. 110, p. 1439, 1996.

[10] S. Bandyopadhyay and V. P. Roychowdhury, "Computational paradigms in nanoelectronics: Quantum coupled single electron logic and neuromorphic networks," *Jap. J. Appl. Phys.*, vol. 35, pt. 1, pp. 3350–3362, 1996.

[11] ——, "Quantum coupled architectures for logic circuits: Issues and problems," in *Proc. Int. Conf. Quantum Dev. Circ.*, Alexandria, Egypt, June 4–8, 1996, K. Ismail, S. Bandyopadhyay, and J.-P. Leburton, Eds. London, U.K.: Imperial College, pp. 271–276.

[12] R. Landauer, "Is quantum mechanics useful?," *Phil. Trans. Royal Soc. London A*, vol. 353, pp. 367–376, 1995.

[13] ——, "Need for critical assessment," *IEEE Trans. Elec. Dev.*, vol. 43, pp. 1637–1639, 1996.

[14] V. P. Roychowdhury, D. B. Janes, S. Bandyopadhyay, and X. Wang, "Collective computational activity in self-assembled arrays of quantum dots: A novel neuromorphic architecture for nanoelectronics," *IEEE Trans. Elec. Dev.*, vol. 43, pp. 1688–1699, Oct. 1996.

[15] M. P. Anantram and V. P. Roychowdhury, "Metastable states in a linear array of quantum dots," in *Proc. PHYSCOMP'96*, Boston, Nov. 1996.

[16] J. C. Lusth, B. Dixon, D. J. Jackson, and S. L. Burkett, "Quantum dot cellular automata and the problem of unbalanced logic gates," in *Proc. Int. Conf. Quantum Dev. Circ.*, Alexandria, Egypt, June 4–8, 1996, K. Ismail, S. Bandyopadhyay, and J.-P. Leburton, Eds. London, U.K.: Imperial College, pp. 289–297.

[17] P. Balasingam and V. P. Roychowdhury, "Nanoelectronic functional devices," Tech. Rep. TR-EE 94-24, Sch. Elec. Eng., Purdue Univ., West Lafayette, IN, July 1994.

[18] D. W. Jordan and P. Smith, *Nonlinear Ordinary Differential Equations*. Oxford: Clarendon, 1987.

[19] A. A. Andronov, A. A. Vitt, and S. E. Khaikin, *Theory of Oscillators*. New York: Dover, 1966.

[20] R. P. Andres, J. D. Bielefeld, J. I. Henderson, D. B. Janes, V. R. Kolagunta, C. P. Kubiak, W. J. Mahoney, and R. G. Osifchin, "Self-assembly of a two-dimensional superlattice of molecularly linked metal clusters," *Sci.*, vol. 273, pp. 1690–1693, Sept. 20, 1996.

[21] D. B. Janes, V. R. Kolagunta, R. G. Osifchin, J. D. Bielefeld, R. P. Andres, J. I. Henderson, and C. P. Kubiak, "Electronic conduction through two-dimensional arrays of nanometer diameter metal clusters," *Superlattices and Microstructures*, vol. 18, pp. 275–282, 1995.

[22] N. S. Bakhvalov, G. S. Kazachca, K. K. Likharev, and S. I. Serdyukova, "Single-electron solitons in two-dimensional tunnel structures," *Sov. Phys. JETP*, vol. 68, no. 3, pp. 581–587, Mar. 1989.

[23] U. Geigenmüller and G. Schön, "Single-electron effects in arrays of normal tunnel junctions," *Europhys. Lett.*, vol. 10, pp. 765–770, Dec. 1989.

[24] D. K. Schroder, "Advanced MOS devices," in *Modular Series on Solid State Devices*, G. W. Neudeck and R. F. Pierret, Eds. Reading, MA: Addison-Wesley, 1987.

[25] Semiconductor Industry Assoc., *The National Technology Roadmap for Semiconductors*, 1994.

[26] D. B. Tuckerman and R. F. W. Pease, *IEEE Elec. Dev. Lett.*, vol. 2, p. 126, 1981.
[27] S. Bandyopadhyay, A. E. Miller, H.-C. Chang, G. Banerjee, V. Yuzhakov, D.-F. Yue, R. E. Ricker. S. Jones, J. A. Eastman, E. Baugher, and M. Chandrasekhar, "Electrochemically assembled quasiperiodic quantum dot arrays," *Nanotechnology*, vol. 7, pp. 360–371, 1996.
[28] B. W. Alphenaar, Z. A. K. Durrani, A. P. Heberle, and M. Wagner, "Resistance bistability in resonant tunneling diode pillar arrays," *Appl. Phys. Lett.*, vol. 66, pp. 1234–1236, 1995.
[29] A. E. Miller, S. Bandyopadhyay, G. Banerjee, D.-F. Yue, and M. Chandrasekhar, unpublished, 1995; S. Bandyopadhyay, A. E. Miller, and M. Chandrasekhar, "Optical, electronic, magnetic, and superconducting properties of quasiperiodic quantum dot arrays synthesized by a novel electrochemical technique," in *Proc. SPIE, Photonics West '95,* San Jose, CA, Feb. 1995, vol. 2397, pp. 11–30.
[30] R. E. Ricker, A. E. Miller, D.-F. Yue, G. Banerjee, and S. Bandyopadhyay, "Nanofabrication of a quantum dot array: Atomic force microscopy of electropolished aluminum," *J. Elec. Mat.*, vol. 25, pp. 1585–1592, 1996.
[31] J. M. Tour and J. S. Schumm, "Synthesis of molecular wires by a novel divergent/convergent approach," *Polym. Prepr.*, vol. 34, pp. 368–369, 1993.
[32] J. I. Henderson, S. Feng, G. M. Ferrence, T. Bein, and C. P. Kubiak, "Self-assembled monolayers of dithiols, diisocyanides and isocyanothiols on gold: 'Chemically sticky' surfaces for covalent attachment of metal clusters and studies of interfacial electron transfer," *Inorg. Chim. Acta,* vol. 242, pp. 115–124, 1996.
[33] M. P. Samanta, W. Tian, S. Datta, J. I. Henderson, and C. P. Kubiak, "Electronic conduction through organic molecules," *Phys. Rev. B,* vol. 53, p. R7626, 1996.
[34] R. P. Andres, T. Bein, M. Dorogi, S. Feng, J. I. Henderson, C. P. Kubiak, W. J. Mahoney, R. G. Osifchin, and R. Reifenberger, "Coulomb staircase at room temperature in a self-assembled molecular nanostructure," *Sci.*, vol. 272, p. 1323, 1996.
[35] S. A. Blanton, A. Dehestani, P. C. Lin, and P. Guyot–Sionnest, "Photoluminescence of single semiconductor nanocrystallite by two-photon excitation spectroscopy," *Chem. Phys. Lett.*, vol. 229, p. 317, 1994; M. E. Schmidt, S. A. Blanton, M. A. Hines, and P. Guyot–Sionnest, "Size-dependent two-photon excitation spectroscopy of CdSe nanocrystals," *Phys. Rev. B,* vol. 53, p. 12629, 1996.
[36] M. P. Patkar, T. P. Chin, J. M. Woodall, M. S. Lundstrom, and M. R. Melloch, "Very low resistance nonalloyed contacts using low-temperature molecular beam epitaxy of GaAs," *Appl. Phys. Lett.*, vol. 66, p. 1412, 1995.
[37] T.-B. Ng, D. B. Janes, D. McInturff, and J. M. Woodall, "Inhibited oxidation in low-temperature grown GaAs surface layers observed by photoelectron spectroscopy," *Appl. Phys. Lett.*, vol. 69, pp. 3551–3553, 1996.
[38] S. Hong, D. B. Janes, D. McInturff, R. Reifenberger, and J. M. Woodall, "Stability of a low-temperature grown GaAs surface layer following air exposure using tunneling microscopy," *Appl. Phys. Lett.*, vol. 68, pp. 2258–2260, Apr. 15, 1996.
[39] R. G. Nuzzo and D. L. Allara, "Adsorption of bifunctional organic disulfides on gold surfaces," *J. Amer. Chem. Soc.*, vol. 105, pp. 4481–4483, 1983.
[40] P. E. Laibinis, R. G. Nuzzo, and G. M. Whitesides, "Structure of monolayers formed by coadsorption of two n-alkanethiols of different chain lengths on gold and its relation to wetting," *J. Phys. Chem.*, vol. 96, p. 5097, 1992.
[41] A. Kumar and G. M. Whitesides, "Patterned condensation figures as optical diffraction gratings," *Sci.*, vol. 263, pp. 60–62, 1994.
[42] M. J. Lercel, R. C. Tiberio, P. F. Chapman, H. G. Craighead, C. W. Sheen, A. N. Parikh, and D. L. Allara, "Self-assembled monolayer electron-beam resists on GaAs and $SiO_2$," *J. Vac. Sci. Tech. B,* vol. 11, pp. 2823–2828, 1993.
[43] S. Y. Chou, P. R. Krauss, and P. J. Renstrom, "Imprint lithography with 25-nanometer resolution," *Sci.*, vol. 272, pp. 85–87, 1996.
[44] T. Sato, D. G. Hasko, and H. Ahmed, "Patterned deposition of colloidal particle monolayer," *Symp. Single Elec. Nanoelectronic, 190th Meet. Electrochem. Soc.*, San Antonio, TX, Oct. 6–11, 1996.
[45] A. Aviram and M. A. Ratner, "Molecular rectifiers," *Chem. Phys. Lett.*, vol. 29, pp. 277–283, 1974.
[46] V. Mujica, M. Kemp, A. Roitberg, and M. Ratner, "Current–voltage characteristics of molecular wires: Eigenvalue staircase, Coulomb blockade and rectification," *J. Chem. Phys.*, vol. 104, pp. 7296–7305, 1996.
[47] C. M. Fischer, M. Burghard, S. Roth, and K. von Klitzing, "Organic quantum wells: Molecular rectification and single electron tunneling," *Europhys. Lett.*, vol. 28, p. 129, 1994.
[48] J. R. Sambles and A. S. Martin, "Molecular rectification," *Phys. Scripta,* vol. T-49, pp. 718–720, 1993.
[49] P. A. Dowben, private communication.
[50] J. M. Carpinelli, E. W. Plummer, D. Byun, and P. A. Dowben, "Scanning tunneling microscopy study of intermediates in the dissociative adsorption of *Closo*-1, 2 dicarbadodecaborane on Si(111)," *J. Vac. Sci. Tech.*, vol. B-13, pp. 1203–1206, 1995.
[51] E. E. Ehrichs, S. Yoon, and A. L. deLozanne, "Direct writing of 10 nm features with the scanning tunneling microscope" *Appl. Phys. Lett.*, vol. 53, p. 2287, 1988; E. E. Ehrichs, R. M. Silver, and A. L. deLozanne, "Direct writing with scanning tunneling microscope," *J. Vac. Sci. Tech.*, vol. A-6, p. 540, 1988.
[52] A. D. Kent, T. M. Shaw, S. von Molnar, and D. D. Awschalom, "Growth of high aspect ratio nanometer scale magnets with chemical vapor deposition and scanning tunneling microscopy," *Sci.*, vol. 262, p. 1249, 1993.
[53] E. E. Ehrichs and A. L. deLozanne, *Nanostructure Physics and Fabrication,* M. A. Reed and W. P. Kirk, Eds. New York: Academic, 1989, p. 441.
[54] D. S. Saulys, A. Ermakov, E. L. Garfunkel, and P. A. Dowben, "Electron-beam-induced patterned deposition of allylcyclopentadienyl palladium using scanning tunneling microscopy," *J. Appl. Phys.*, vol. 76, p. 7639, 1994.
[55] G. Dujardin, R. E. Walkup, and P. Avouris, "Dissociation of individual molecules with electrons from the tip of a scanning tunneling microscope," *Sci.*, vol. 255, p. 1232, 1992.
[56] T. Urisu and H. Kyuragi, "Synchrotron-radiation-excited chemical vapor deposition and etching," *J. Vac. Sci. Tech.*, vol. B-5, p. 1436, 1987.
[57] F. K. Perkins, M. Onellion, S. Lee, D. Li, J. Mazuroski, and P. A. Dowben, "Synchrotron radiation induced deposition of boron and boron carbide films from boranes and carboranes II; Nido-2, 3-diethyl-2, 3-dicarbahexaborane," *Appl. Phys.*, vol. A-54, p. 442, 1993.
[58] D. Byun, S.-D. Hwang, P. A. Dowben, F. K. Perkins, F. Flipps, and N. J. Ianno, "Heterojunction fabrication by selective chemical vapor deposition induced by synchrotron radiation," *Appl. Phys. Lett.*, vol. 64, pp. 1968–1970, 1996.

**Vwani P. Roychowdhury** (Member, IEEE) received the B.Tech. degree from the Indian Institute of Technology, Kanpur, and the Ph.D. degree from Stanford University, Stanford, CA, in 1982 and 1989, respectively, both in electrical engineering.

From 1989 to 1991, he was a Research Associate with the Department of Electrical Engineering at Stanford University. From 1991 to 1994, he was an Assistant Professor of Electrical Engineering at Purdue University, West Lafayette, IN, and became Associate Professor in 1994. Currently, he is an Associate Professor of Electrical Engineering, University of California, Los Angeles, CA. His research interests include parallel algorithms and architectures, design and analysis of neural networks, application of computational principles to nanoelectronics, special purpose computing arrays, VLSI design, and fault-tolerant computation. He has co-authored several books including *Discrete Neural Computation: A Theoretical Foundation* (Prentice-Hall, 1995) and *Theoretical Advances in Neural Computation and Learning* (Kluwer, 1994).

Dr. Roychowdhury was named the General Motors Faculty Fellow in the School of Electrical Engineering at Purdue University from 1992 to 1994. For his outstanding performance in the classroom, he received the 1993 Ruth and Joel Spira Outstanding Teacher Award.

**David B. Janes** (Member, IEEE) received the B.A. degree in physics (summa cum laude) from Augustana College, Rock Island, IL, in 1980 and the B.S.E.E (highest honors), M.S.E.E., and the Ph.D. degrees from the University of Illinois, Urbana-Champaign, IL, in 1980, 1981, and 1989, respectively. His doctoral research focused on the effect of deep levels on the performance of acoustic charge transport devices.

From 1981 to 1985, he was a Research Scientist in the Research Division of Raytheon Company, where he worked on microwave semiconductor devices, including GaAs IMPATT diodes and MESFET's and monolithic microwave integrated circuits. From 1981 to 1996, he was an Assistant Professor at the School of Electrical Engineering, Purdue University, and currently is a research engineer at Purdue University, West Lafayette, IN. He is presently engaged in experimental research on mesoscopic semiconductor devices and metallic cluster arrays, characterization of novel semiconductor heterostructures, and microwave devices based on compound semiconductor heterostructures.

**Supriyo Bandyopadhyay** (Senior Member, IEEE) received the B.Tech. degree in electronics and electrical communications engineering from the Indian Institute of Technology, Kharagpur, in 1980, the M.S. from Southern Illinois University, Carbondale, IL, and the Ph.D. from Purdue University, West Lafayette, IN, both in electrical engineering, in 1982 and 1985, respectively.

During 1986 he was Visiting Assistant Professor at the School of Electrical Engineering, Purdue University, West Lafayette, IN. From 1986 to 1996, was a faculty member at the University of Notre Dame, Notre Dame, IN. Currently, he is Professor of Electrical Engineering at University of Nebraska, Lincoln, NE. He has authored more than 180 articles related to quantum devices, hot electron transport, superlattice physics, nonlinear optics, and superconductivity. He has chaired many international conferences and edited books in these fields, and serves on the editorial board of *Physics of Low Dimensional Systems* (Vasov, Moscow, Russia).

Dr. Bandyopadhyay is a member of Tau Beta Pi, Eta Kappa Nu, Sigma Pi Sigma, and the Electrochemical Society. He is listed in *Who's Who in Science and Technology*, the *Dictionary of International Biography* (Cambridge, U.K.), and *Sterling's Who's Who*.

Nanotechnology 7 (1996) 360–371. Printed in the UK

# Electrochemically assembled quasi-periodic quantum dot arrays

**S Bandyopadhyay†, A E Miller‡, H C Chang‡, G Banerjee‡, V Yuzhakov‡, D-F Yue§, R E Ricker‖, S Jones‖, J A Eastman¶, E Baugher+ and M Chandrasekhar+**

† Department of Electrical Engineering, University of Nebraska, Lincoln, Nebraska 68588-0511, USA
‡ Department of Chemical Engineering, University of Notre Dame, Notre Dame, Indiana 46566, USA
§ SGS-Thomson Microelectronics, Inc., Carrolton, Texas 75006, USA
‖ Materials Science and Engineering Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA
¶ Argonne National Laboratory, Argonne, Illinois 60439, USA
+ Department of Physics, University of Missouri-Columbia, Columbia, Missouri 65211, USA

**Abstract.** We describe two electrochemical self-assembly processes for producing highly ordered quasi-periodic arrays of quantum dots on a surface. The advantages of these techniques are: (i) they are 'gentle' and do not cause radiation damage to nanostructures unlike beam lithography, (ii) they have high throughput and are amenable to mass production unlike direct-write lithography, (iii) structures can be delineated on non-planar substrates, and (iv) the techniques are potentially orders of magnitude cheaper to implement than conventional nanosynthesis. Samples produced by these techniques have been characterized by microscopy, optical and transport measurements, Auger and x-ray. These measurements reveal intriguing properties of the nanostructures. In this paper, we describe our initial results and show the promise of such techniques for low-cost and high-yield nanosynthesis.

## 1. Introduction

It is well known that conventional fine line lithography (electron beam, focused ion beam, x-ray, STM/AFM etc.) has at least one of two major shortcomings: (i) process related damages incurred through exposure to high-energy beams during pattern delineation (writing) [1–4] or through reactive ion etching (pattern transfer) [5], and (ii) slow throughput associated with direct-write lithography where each wafer is patterned *serially* one at a time. Recently, parallel electron-beam columns and arrayed nanoprobes have been implemented to increase the throughput and introduce some parallelism into the process, but the cost of such systems is usually prohibitive [6].

In an effort to mitigate these problems, we are developing two electrochemical self-assembly techniques for nanosynthesis that fine-tune a current 'low technology' process routinely available in the anodizing industry into a 'high technology' process for mass-fabricating dense ($>10^{12}$ cm$^{-2}$) two-dimensional *quasi-periodic* arrays of metallic, semiconducting or superconducting quantum dots. Samples produced by these techniques have been directly imaged with transmission electron microscopy (TEM), field emission scanning electron microscopy (FESEM) and atomic force microscopy (AFM). Quantitative image analysis carried out by an image capture and digitization system indicates that the average diameter of dots produced by one of these techniques can be as small as 10 nm with a 1 nm variation across a wafer, while the average interdot separation can be as small as 40 nm with a 2 nm variation. This tight size control is comparable with that obtained with advanced lithographic techniques [7], or molecular-beam epitaxy (MBE) growth of randomly distributed self-assembled quantum dots realized through coherent islanding in Stranski–Krastanow mode [8]. The quantum dots produced by the electrochemical technique have also been characterized by energy-dispersive analysis of x-ray, AC susceptometry, photoluminescence, Raman spectroscopy, absorption and reflectance studies, ellipsometry, and transport studies to investigate their magnetic, optical, electronic and superconducting properties. In section 4, we describe some of these results; but first we describe the fabrication processes.

## 2. Self-assembly of a mask for a periodic array of quantum dots: electropolishing of aluminum

In this section, we describe the first of the two electrochemical processes that creates a self-assembled *mask* for a *highly periodic* array of quantum dots. The array

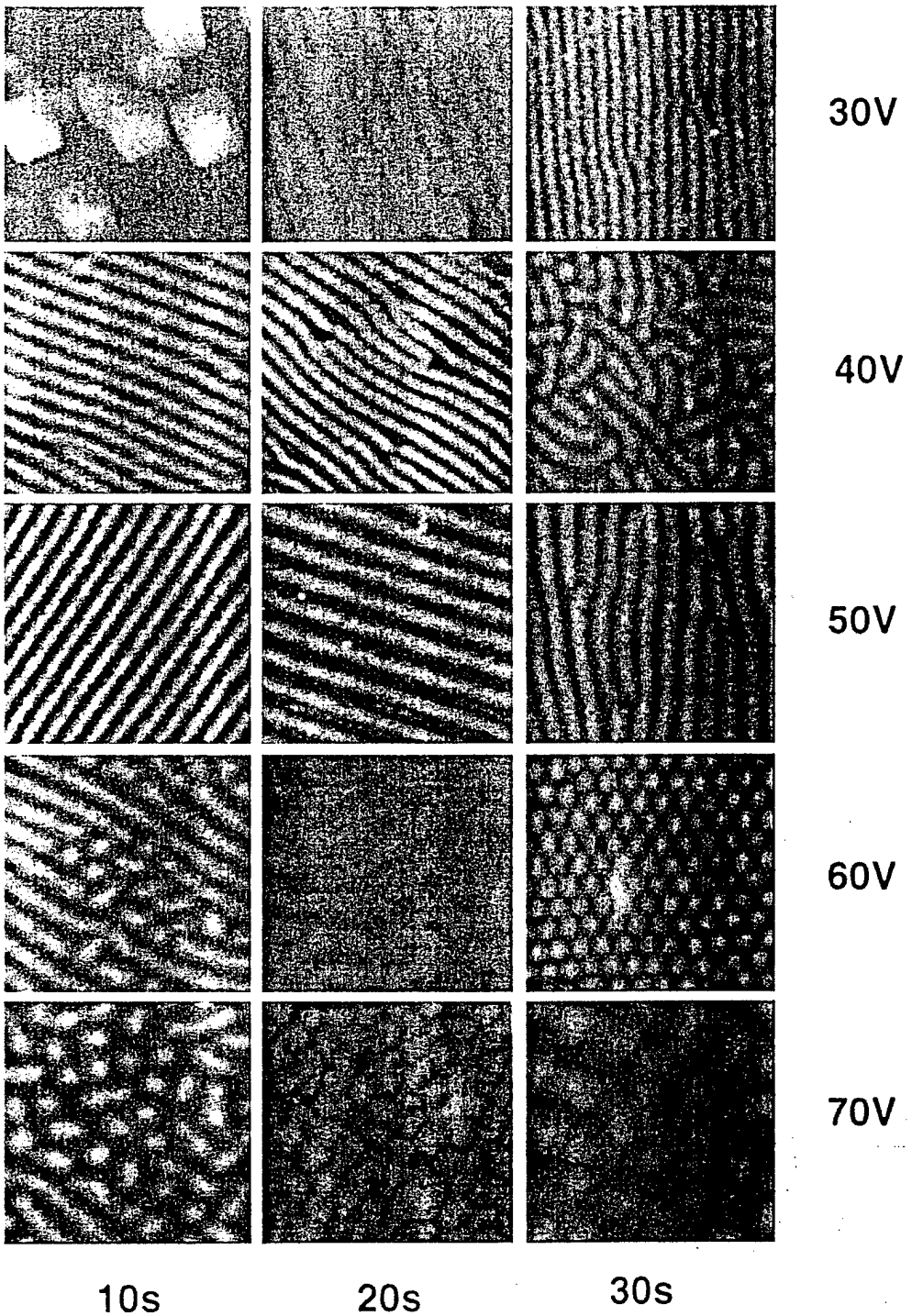



**Figure 1.** Atomic force micrographs of various patterns that are formed on the surface of electropolished aluminum films after electropolishing in perchloric acid, butyl cellusolve, ethanol and water for various durations of time and at various voltages.

has a hexagonal order (a two-dimensional Fourier transform of an atomic force micrograph of the pattern shows six disinct peaks attesting to the high degree of regimentation and the hexagonal order).

To create this structure, we start with a 99.99% pure Al film which could be evaporated on any chosen conducting substrate such as doped silicon. This film is then electropolished in a solution of 62 cc perchloric acid, 700 cc ethanol, 100 cc butyl cellusolve and 137 cc distilled water for various time durations and at various voltages. The electropolishing process patterns the surface of the aluminum film into a dimpled and undulating landscape which may or may not have spatial ordering depending on the duration and voltage of electropolishing. Atomic force micrographs (raw data) are shown in figure 1 indicating the patterns that form at various voltages and for various durations of electropolishing. The patterns evolve from a random geography through highly ordered stripes (50 V, 10 s) and an 'egg-carton' pattern (60 V, 30 s). Obviously, the stripes and the egg-carton are

# TWO TYPICAL ELECTROPOLISHED SURFACE STRUCTURES

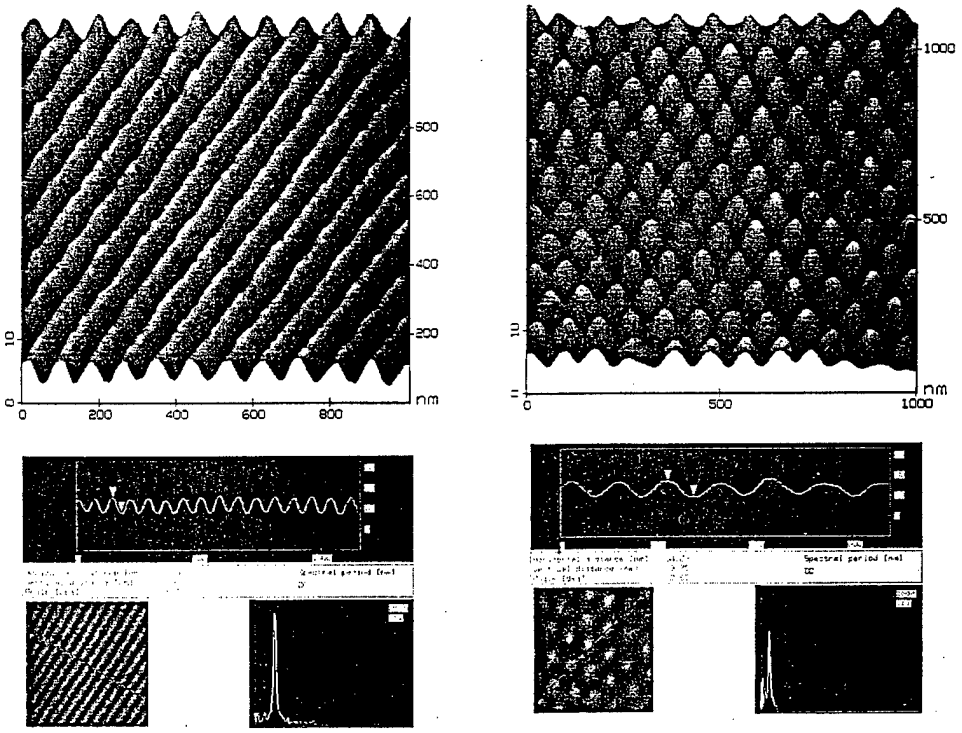



**Figure 2.** Magnified atomic force micrographs (raw data) of stripes and 'egg-carton' patterns that form on the surface of electropolished aluminum. Stripes form after electropolishing at 50 V for 10 s and egg-cartons form after electropolishing at 60 V for 30 s. These patterns can be used as self-assembled masks for quantum wires and dots. The figure also shows depth profiles along the lines indicated. Well resolved peaks in the Fourier transform of these profiles attest to the high degree of periodicity and order within the scanned area.

useful for making quantum wires and dots respectively. Figure 2 shows a magnified micrograph of the stripe and dot patterns with depth profiles along arbitrary lines drawn on the surface. Fourier transforms of the depth profiles show very narrow and well resolved peaks attesting to the high degree of periodicity of the structures. These structures are incomparably more periodic and ordered than those produced by coherent islanding in Stranski–Krastanow growth [8]. The 'pitch' or 'period' of the pattern is about 100 nm (peak to peak separation) and the height of a crest above a trough is about 3 nm. The hills appear to be conical in shape with a full width at half maximum of about 30 nm.

These structures can be used as self-assembled masks for producing highly periodic arrays of quantum wires and dots. This is achieved as follows (see figure 3). A thin film of aluminum (about 10 nm thick) is evaporated on a chosen substrate. After electropolishing to produce the desired landscape (stripes or egg-carton), the troughs are selectively etched away in a suitable etchant (note that the peak to valley ratio, or the ratio of the Al film thicknesses at the crests and troughs is about 3:2). The etching leaves behind the crests which form isolated islands of metal dots in a periodic pattern on the surface of the semiconductor. These Al islands can be used as a mask for mesa-etching of quantum dots using reactive ion etching. Finally, a layer of a material can be deposited on the surface to electrically connect the quantum dots, if desired. The four steps in the process are depicted in figure 3. Note that electrical connection between the quantum dots can also be established via molecular-wire linked clusters [9–11] deposited between the dots. The latter is an established technology at this time.

### 2.1. Theory of pattern formation during electropolishing

The formation of the various patterns during electropolishing can be explained by field-assisted dissolution of metals. We are able to model and mathematically capture the proper conditions for the formation of the patterns on the Al surface and the spatio temporal scalings. During electropolishing, a double layer is formed on the liquid electrolyte side near the constant-potential metal surface as ions are attracted to it under the action of the electropolishing voltage. This sets up a potential gradient normal to the metal–electrolyte interface. The electrolyte contains organic molecules like ethanol and butyl cellusolve. The potential gradient reduces the threshold for the desorption of the organic molecule and consequently, crests, with higher poten-

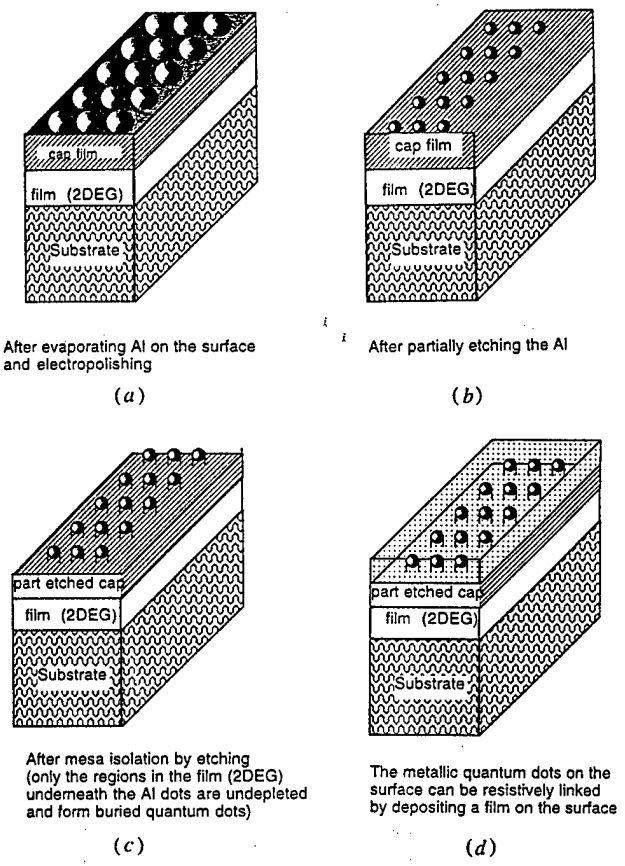



**Figure 3.** Figure indicating the four steps in using the electropolished aluminum as a self-assembled mask to create a periodic array of electrically linked quantum dots on a surface. Such structures have applications in nanoelectronic computing architectures utilizing single electron effects in quantum dots (see [20]). (*a*). Structure after electropolishing a 100 nm Al film deposited on a semiconductor layer, (*b*) after etching the troughs selectively to leave behind the crests that form isolated islands on the surface, (*c*) shallow mesa etching with reactive ions using the aluminum dots as a mask, (*d*) electrically connecting the dots by depositing a layer between the mesas.

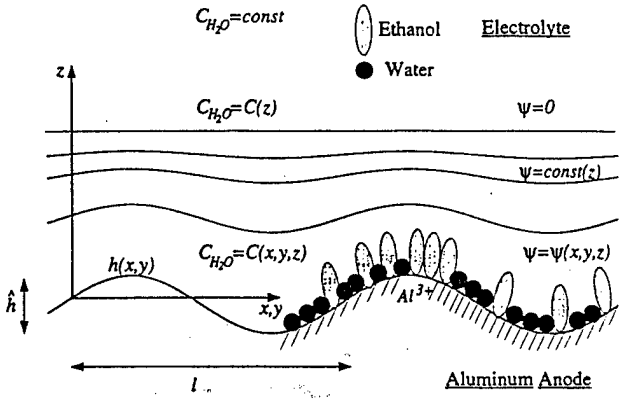



**Figure 4.** Figure to explain the chemistry of the formation of various surface morphologies on electropolished aluminum. Regions covered by ethanol molecules are protected from dissolution and this results in the formation of various undulating patterns. Ethanol molecules are selectively adsorbed on the crests because the higher potential gradient there reduces the threshold for desorption of ethanol molecules. The constant potential ($\psi$) surfaces are schematically depicted.

tial gradients, are preferentially covered by these molecules. The dissolution of $Al^{3+}$ is then hindered at the crests and troughs dissolve faster (figure 4). This destabilizing mechanism is countered by a curvature stabilizing mechanism which favours surface Al atoms on a concave surface over dissolved $Al^{3+}$ ions. As a result, short wavelength disturbances are damped. More importantly, as troughs dissolve deeper, the stabilizing effect at the valley increases while the destabilizing potential gradient in the double layer decreases. Consequently, a finite equilibrium amplitude $h(x, y)$ of the patterns is reached where the two curvature-dependent mechanisms balance to yield regular patterns. The two-dimensional distribution of the amplitude $h(x, y)$ obeys the generalized nonlinear Kuramoto–Sivashinsky equation. Numerical simulation to extract the two-dimensional profile of $h(x, y)$ shows patterns that are strikingly similar to the experimentally observed patterns. The simulations show clearly that depending on the voltage and duration of electropolishing, one does indeed get stripes or 'egg-carton' pattern or the random patterns as shown in figure 1. Details of these results will be published elsewhere.

We are able to estimate the curvature-dependent threshold for desorption of organic molecules and dissolution of the Al ion by studying the free energy of the pertinent chemical complexes in the presence of a field. This information allows us to predict the critical voltages beyond which patterns appear, the time scales for transition from stripes to dots, and the characteristic length scales of both patterns. Because of the activated desorption process, the wavelength or 'pitch' of the patterns scales as the exponential of the square of the electropolishing voltage as shown in figure 5. We also believe that by reducing the temperature of the electrolyte during electropolishing, we can reduce the pitch of the pattern and also the full-width-half-maximum of the crests which will result in smaller quantum dots and a denser array. This, however, remains to be tested at this time.

## 3. Nanoporous films formed by anodization of aluminum

We now describe the second process for the creation of quasi-periodic arrays of quantum dots. This technique produces quantum dot arrays with somewhat reduced periodicity, but the individual dot diameter is about five times smaller ($\sim$ 10 nm). We DC-anodize the electropolished aluminum film produced in the previous step under a constant current density of 40 mA $cm^{-2}$ at room temperature in 15% $H_2SO_4$ (pH = 0.6). The Al film is used as the anode and a platinum mesh as the cathode. Anodization creates a porous alumina film on the surface of Al with a regimented, hexagonal close-packed arrangement of nanopores as shown in the idealized diagram in figure 6. Formation of this film has been known in the anodizing industry for the last 50 years and has been recently revisited and studied by a number of researchers [12–15]. A bright field transmission electron micrograph of an actual porous alumina film is shown in figure 7. Image analysis of such

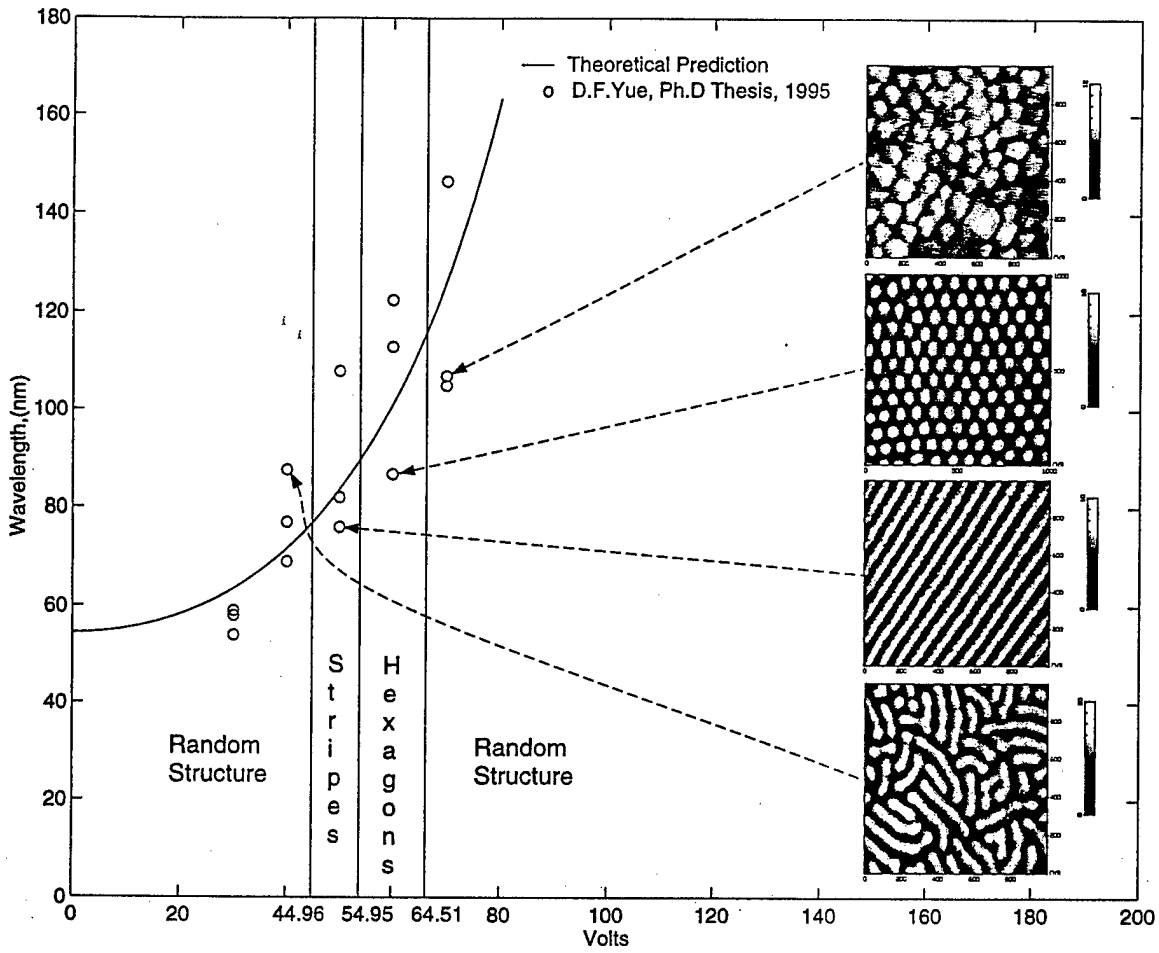



**Figure 5.** Dependence of the wavelength (or pitch) of the periodic patterns on electropolishing voltages. The solid curve is the computed theoretical result and the experimental data are due to one of us (D-F Yue).

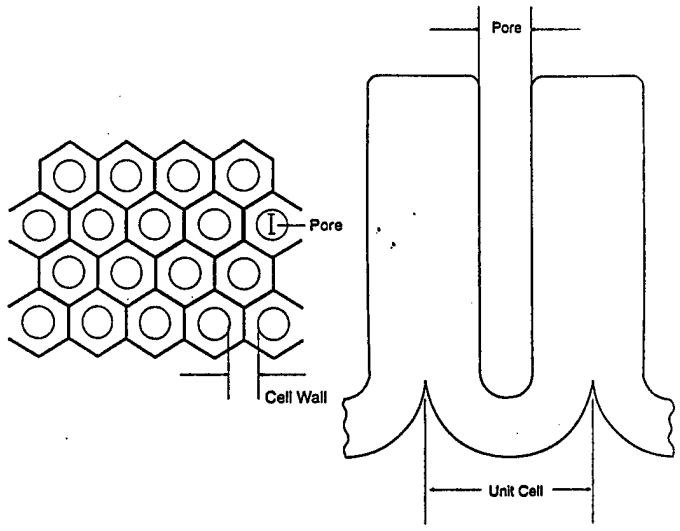



**Figure 6.** Schematic depiction of the nanoporous film that forms on the surface of aluminum upon anodization. The pores have a hexagonal order. Both the top view and the side view of a pore are shown.

micrographs reveals that typically the average diameter of the pores is ~ 10 nm with a 7% standard deviation and the center-to-center separation is about 40 nm with a 2% standard deviation. Histograms of the pore diameter distribution are approximately Gaussian.

Based on extensive TEM, field-emission SEM and AFM characterization of the nanoporous films formed under various conditions of anodization, we have arrived at the following conclusions.

- Other factors being the same, the pore density decreases and the pore diameter increases slightly with increasing anodizing current density.

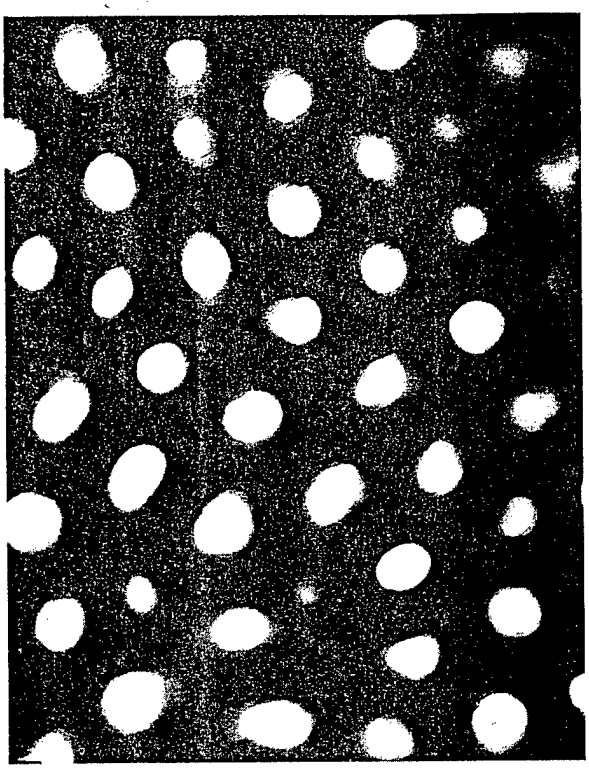


**Figure 7.** A bright field transmission electron micrograph of a nanoporous film formed by anodizing aluminum in 15% sulfuric acid at room temperature with a current density of 40 mA $cm^{-2}$. The average pore diameter is 15 nm. The ordering in some regions is clearly hexagonal. TEM samples are produced by dissolving the aluminum backing from the nanoporous film with bromine/methanol and capturing the released porous film in TEM grids.

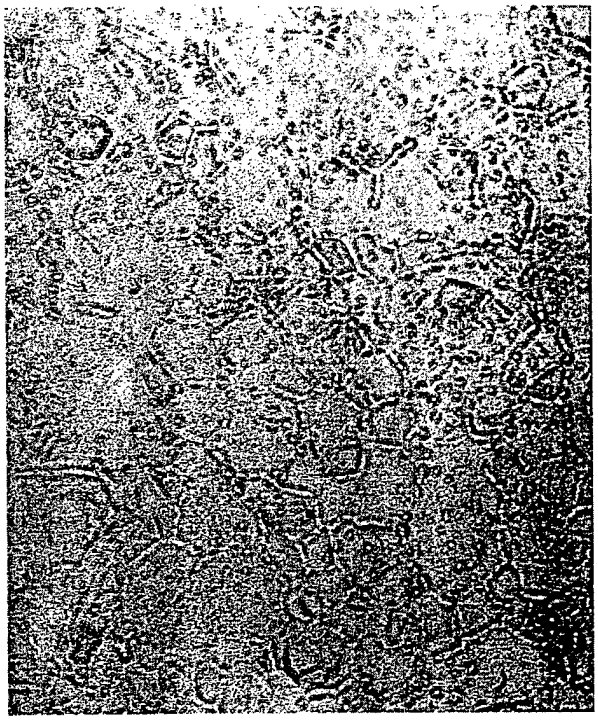


**Figure 8.** A bright field transmission electron micrograph of pores nucleating on a 300 nm thick layer of Al evaporated on a single crystal silicon substrate. The evaporated layer was composed of ultrafine grain aluminum with typical grain diameters of 1 $\mu$m. The anodization was carried out in 15% sulfuric acid with a current density of 30 mA $cm^{-2}$. The magnification is 90 000×. Note that pores begin to nucleate at first along grain boundaries which are regions of increased surface elastic energy.

- The control over pore size, pore density, and most importantly, regimentation (or spatial ordering) improves with increasing current density. Our most regimented pore arrangements are those that are produced at the highest current densities. The highest current densities also result in the minimum standard deviation in the pore diameter, thereby affording the best feature size control.
- Pores tend to nucleate in regions of increased surface elastic energy. Typically grain boundaries are preferred sites as shown in figure 8.
- The pore depth depends strongly on the time of anodization. For the first few seconds of anodization (typically 6–7 s for a current density of 40 mA $cm^{-2}$), a barrier layer of alumina forms on the surface and then the pores begin to nucleate. The pore depth continues to increase with increasing duration of anodization. To make pores that are about 10 nm deep, the time required is usually a few seconds. The pore depth also depends slightly on the underlying crystallographic orientation of the underlying aluminum grain. This is shown clearly in figure 9.

The pore diameter, pore depth, pore density and regimentation can be controlled by fine tuning the temperature, the pH of the acid, the chemical composition of the acid used during anodization (generally $H_2SO_4$ gives the smallest pore diameter while phosphoric and oxalic acids produce much larger pore diameters), and the current density. Pore diameters as small as 5 nm have been reported in the literature [15]. In some situations, quite well regimented arrays of pores can be produced as shown by the transmission electron micrograph in figure 10. The pore densities in such structures can approach $10^{12}$ $cm^{-2}$ which is excellent for applications in high density magnetic storage or high density semiconductor dot arrays for nonlinear optical devices. If the periodicity can be improved slightly, these will also result in self-organized two-dimensional dielectric structures that exhibit a *photonic bandgap* for all electromagnetic wave propagation directions along the plane of the film [16]. The bandgaps will open up at wavelengths corresponding to deep ultraviolet or possibly x-ray.

### 3.1. Theory of pore formation

The mechanism of pore formation is not entirely different from that which undergirds the formation of the undulating patterns during electropolishing. In contrast to electropolishing, where the potential gradient occurs within the double layer on the metal surface, the potential gradient during anodization lies mainly in the oxide layer which is insulating. Dissolution occurs because the surface oxygen atoms prefer to be surrounded by water molecules in the bulk solution owing to the presence of the electric field. Pitting initiates at a thinner part of the oxide layer because

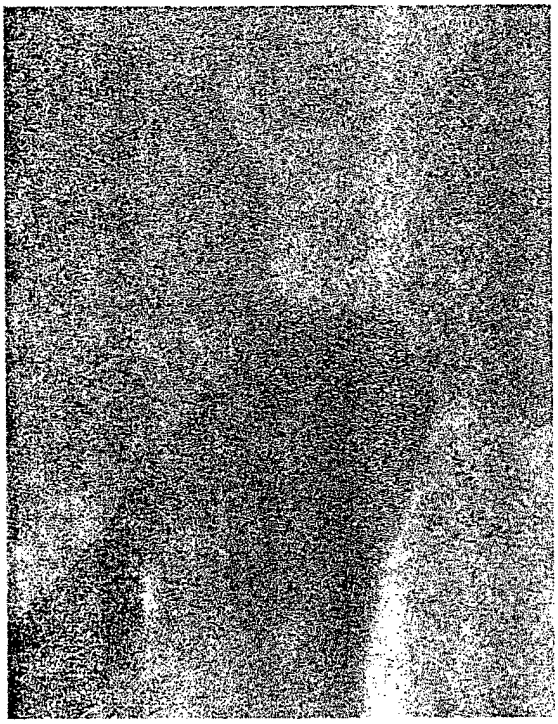


**Figure 9.** A bright field TEM image of pores formed on various grains in a large grained polycrystalline Al substrate anodized in sulfuric acid at room temperature for 6 s with a constant current density of 30 mA $cm^{-2}$. The magnification is 12 000×. Some (but not all) grain boundaries are decorated with extremely regimented arrays of pores indicating once again that regions of increased elastic energy produce better regimentation. Lighter shades indicate smaller film thickness (more transparent to transmission electron beam). Thus, the thickness of the anodized film depends somewhat on the crystallographic orientation of the underlying aluminum.

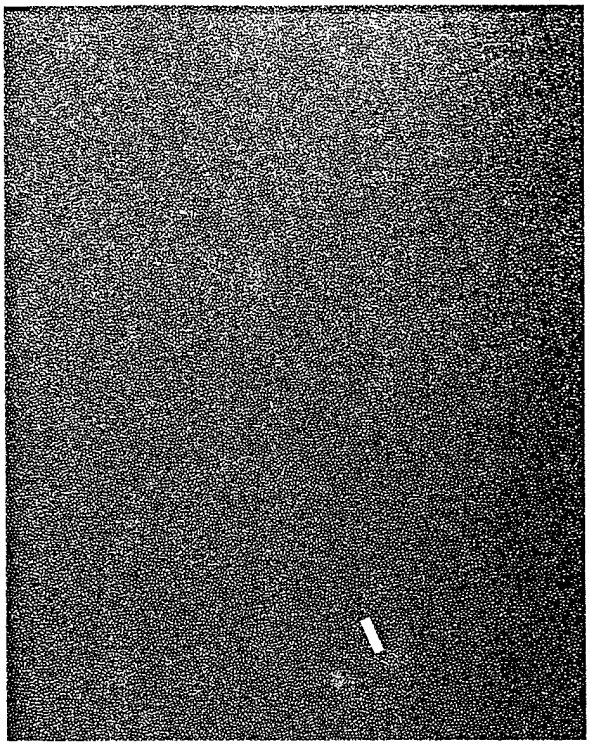


**Figure 10.** A bright field TEM image of pores taken at a low magnification of 30 000×. The white bar is 300 nm. The anodizing current density was 40 mA $cm^{-2}$ and the aluminum film was cold worked. The image shows fairly long range regimentation along strips oriented in the direction of the white bar. These strips are several microns long and few hundreds of nm wide. Adjacent strips are separated by regions of short range regimentation. The pore density is $4 \times 10^{10}$ $cm^{-2}$ which is excellent for high density magnetic storage or semiconductor quantum dot arrays for applications in nonlinear optics.

of the enhanced electric field there. This accelerates the solvation rate of the surface oxygen atoms. Since the potential gradient increases as the oxide layer thins, the pits (or pores) grow self-catalytically because of the positive feedback mechanism. The curvature-reduced dissolution rate serves only to select the size of the pore. Owing to the fact that the potential gradient scales linearly with the oxide layer thickness, the pore separation and the pore diameter scale *linearly* with the anodizing voltage—a drastically different situation from the *exponential* scaling associated with electropolishing. Because the oxide layer dissolution is determined almost solely by its thickness, these pore size scalings are universal correlations independent of the electrolyte composition (see figure 11). The basic mechanism is, however, somewhat complicated by the field driven flux of oxygen ions through the oxide (alumina) film which thickens the layer. The mechanism of pore initiation is schematically depicted in figure 12(*a*). As shown from our numerical construction of the pit/pore in figure 12(*b*), the enhanced potential gradient at a pit also produces a thick oxide layer which obviously stabilizes further growth. This secondary mechanism finally arrests the accelerated pore growth and determines the final pore depth. Unlike the pore diameter and the interpore separation, the scaling of the pore depth with the applied voltage is a function of

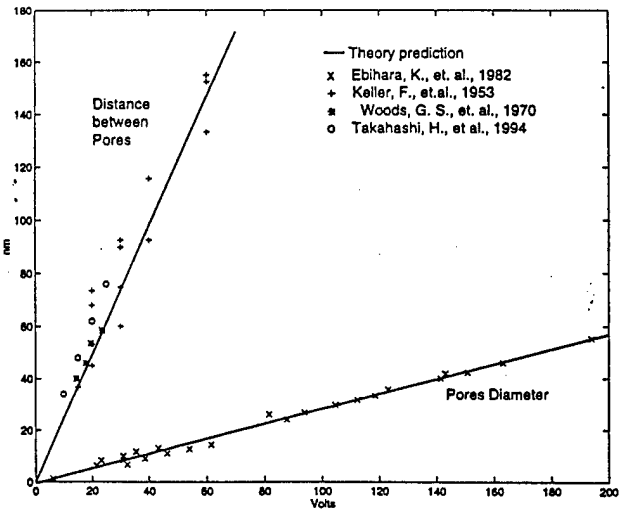



**Figure 11.** Dependence of the pore diameter and interpore separation on the anodizing voltage. The solid lines are the computed theoretical results and the experimental data are those reported by various researchers.

the electrolyte composition, especially the concentration of oxidizing ions.

This model of pore formation should be compared and contrasted with the various models proposed for the

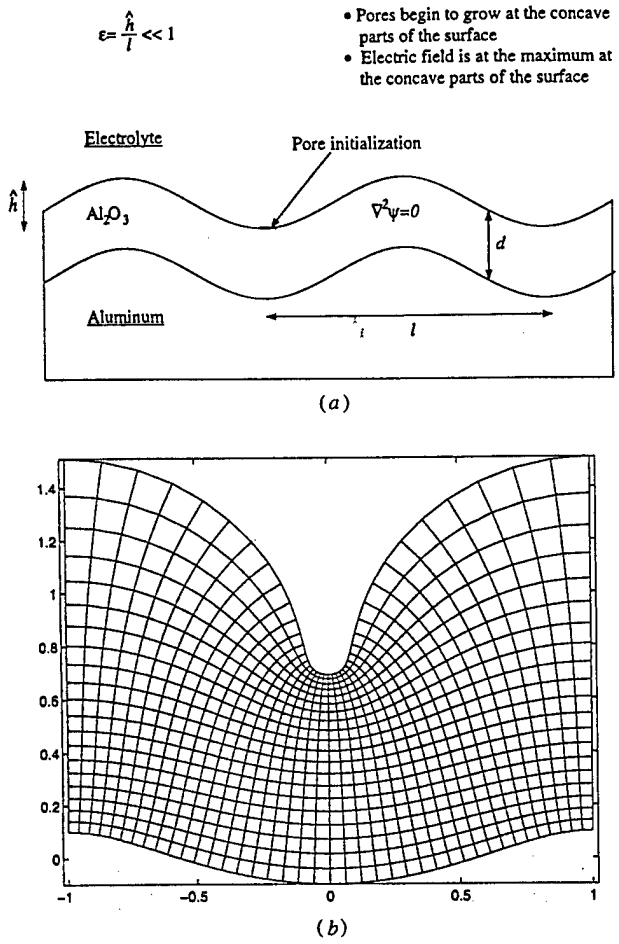



**Figure 12.** (*a*) Schematic representation of the initiation of pore nucleation. (*b*) Numerical construction of a pore showing its formation.

formation of porous silicon [17].

### 3.2. Filling the pores by electrodeposition

To finally create 'quantum dots', the nanopores formed by anodization are selectively filled with the material of interest by electrodeposition. When this is complete, one obtains a two-dimensional quasi-periodic arrangement of quantum dots (or short, vertical quantum wires if the pore depth significantly exceeds the pore diameter) surrounded by $Al_2O_3$ (a dielectric insulator). The $Al_2O_3$ insulator can be removed, if desired, by first depositing a suitable film on the filled pores (third step) and then dissolving away the insulator in an etching solution (final step). This produces a quasi-periodic quantum dot array with the film deposited in the third step acting as the substrate (see figure 13). The last two steps may be unnecessary in most electronic, magnetic and optical applications since the presence of an insulating dielectric surrounding the quantum dots does not interfere with these applications. We describe the basic pore-filling step below.

The pores are filled with the material of interest by electrodeposition. For this, the porous film is first washed in distilled water and dried. The pores are enlarged and 'rounded', if needed, by dissolving part of the surrounding

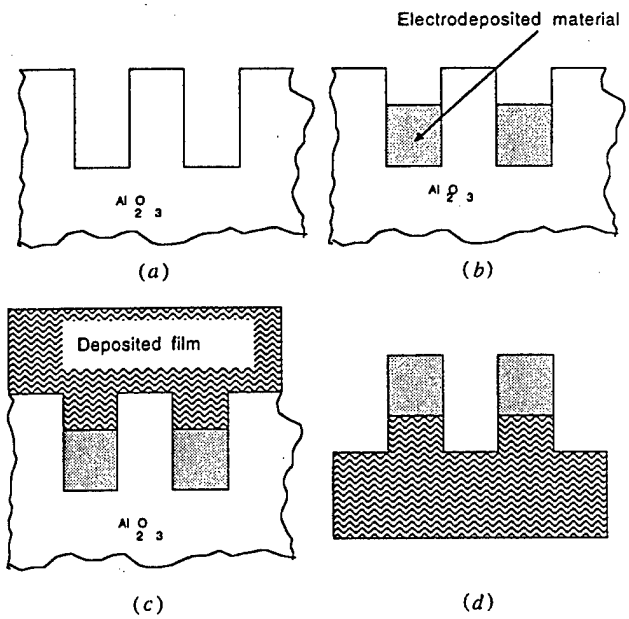
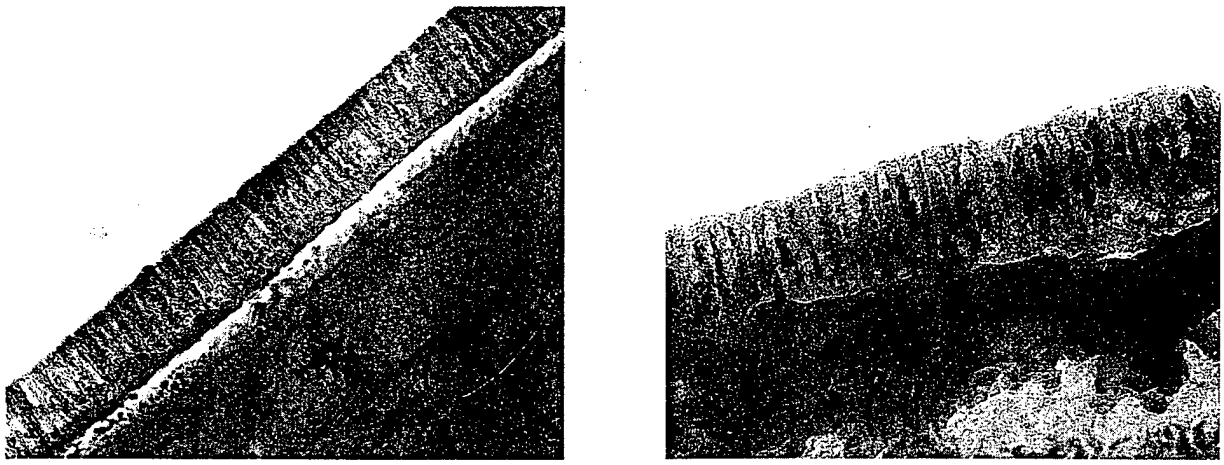



**Figure 13.** Various steps in producing quantum dots by electrodepositing materials of interest within pores. (*a*) Side view of the nanoporous film, (*b*) after electrodepositing the material of interest within the pores, (*c*) depositing a film over the structure, (*d*) selectively etching away the alumina to create mesa quantum dots on the film deposited in the third step acting as the final substrate.

alumina walls in an aqueous acid solution. After this, the film is washed to remove debris and dried.

For electrodepositing most metals (e.g. Fe, Ni, Pb, etc.), the porous film is placed in an acidic, non-cyanide bath that does not attack alumina, and AC current is passed at a suitable rms voltage between 18 and 20 V using the bottom Al substrate and a graphite plate immersed in the bath (or a platinum mesh) as the electrodes. The bath contains a soluble salt of the metal which ionizes to produce the metal ion which is preferentially electrodeposited within the pores since the pores offer the least impedance paths between the two electrodes. The time of electrodeposition determines the depth of pore filling. To control the filling factor (how much of the pore is filled up), one can monitor the surface resistivity of the film or carry out real-time impedance spectroscopy. The surface resistivity correlates very well with the filling factor and can be used to calibrate this process. We show a cross-sectional transmission electron micrograph of nickel filled pores in figure 14. These TEM specimens were produced by ultramicrotomy. For a 5 s electrodeposition, only a small part of the pores is filled resulting in 10 nm sized quantum dots of nickel situated at the bottom of the pores (figure 14(*a*)). For a 20 s electrodeposition, the pores are filled up almost to the brim resulting in short and vertical quantum wires of nickel (200 nm height) within the pores (figure 14(*b*)).

For electrodepositing compound semiconductors like CdS or ZnS within the pores, the alumina film is first immersed in $H_2SO_4$ and AC electrolysis is performed for 10 s with a rms current density of 40 mA $cm^{-2}$ which leaves behind the $S^{2-}$ ion in the pores. This is followed by immersing the film in boiling deionized water containing a 10% solution of $CdSO_4$ or $ZnSO_4$. This solution must be



(*a*) (*b*)

**Figure 14.** A cross-section TEM image of nickel filled pores. (*a*) After electrodepositing nickel for 5 s, and (*b*) after electrodepositing nickel for 20 s. In the first case, we have obtained spherical nickel quantum dots of 10 nm diameter and in the second case, we have obtained cylindrical short quantum wires with 10 nm diameter and 200 nm height.

chemically pure. Quite often, this is the major source of contamination. The cation $Cd^{2+}$ or $Zn^{2+}$ in the solution reacts with the $S^{2-}$ in the pores to form CdS or ZnS inside the pores. The time of immersion determines the amount of compound produced by this reaction and therefore determines the depth of pore filling. Usually not all the $S^{2-}$ reacts with the metal ion so that some excess $S^{2-}$ is left behind in the pores. This does not cause a problem. Deposition of CdS or ZnS takes place within the pores at a slow rate thereby affording excellent control. The bath can be used for several successive operations if material purity is not a serious concern. The water for rinsing is treated by usual chemical treatment or ion exchange before being drained.

The materials electrodeposited within the pores may not be crystalline. However, crystallinity may be restored, at least partially, by thermal annealing following the deposition. For CdS, we anneal at 150 °C for 6–24 h.

The above process can be used to create quantum dots of ZnTe, CdTe, ZnSe, CdTe, GaAs, InAs, InP, GaP, etc. For tellurides, one uses telluric acid instead of sulfuric acid; for arsenides, one uses arsenic acid, for phosphides, one uses phosphoric acid and for selenides, one uses selenic acid.

After the electrodeposition step, one is left with a quasi-periodic arrangement of quantum dots surrounded by $Al_2O_3$. The oxide can be removed, if desired, by following the last two steps illustrated in figure 13. Usually however, this is unnecessary.

Before concluding this section, we would like to point out a few advantages of this technique for producing quasi-periodic quantum dot arrays:

(1) It is a 'gentle' electrochemical technique and does not subject the nanostructures to high-energy beams such as x-ray, electron or ion beam. There is no etching step which would have caused material damage. Consequently, it may result in better material quality when perfected.
(2) It is amenable to mass production. The number of wafers that can be handled simultaneously is limited by the size of the anodizing or electrodeposition bath. For a sufficiently large bath, this number can be several millions.
(3) One can produce nanostructures on curved substrates (such as the surface of a cylinder or sphere) rather than plane substrates exclusively. Anodization of curved substrates and the subsequent ac electrodepostion is no different than for plane surfaces.
(4) The technique is much cheaper than most other techniques for nanosynthesis. The capital investment cost is a few thousand dollars rather than a few million dollars.

## 4. Characterization

We now present some results that characterize the properties of the quantum dots. We chose six different methods for studying the dots.

(1) Analysis of energy dispersive x-ray;
(2) optical characterization such as photoluminescence, absorption and reflectance, ellipsometry and Raman spectroscopy;
(3) magnetic and transport studies;
(4) cross-sectional TEM;
(5) AC susceptometry;
(6) Scanning Auger spectroscopy.

We will discuss some of them here.

### 4.1. Energy-dispersive analysis of x-ray scattering

Energy-dispersive analysis confirms the presence of the material of interest within the pores. The data for Pb quantum dots produced by this technique are shown in figure 15.

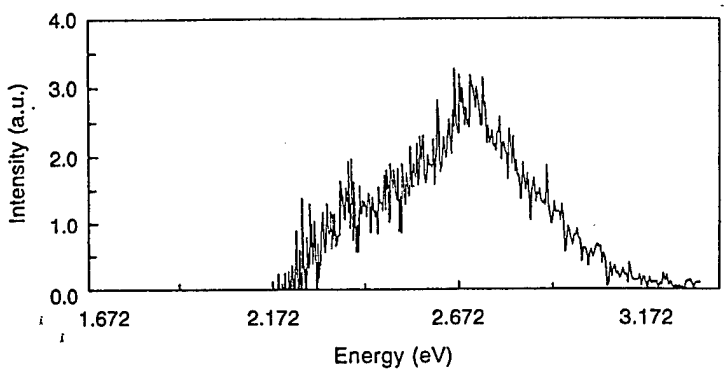



**Figure 16.** Photoluminescence spectrum of CdS quantum dots prepared by electrodepositing CdS within the pores. The peak shows a blue shift of 200 meV from the bulk bandedge indicating quantum confinement. The calculated dot diameter is 6 nm which is less than the TEM observed pore diameter of 13 nm. The difference of 7 nm can be explained by side depletion by interface states, or by the fact that the blue shift is not entirely due to quantum confinement, but partly due to strain. The large full-width-at-half-maximum could be due to lack of complete crystallinity, homogeneous and inhomogeneous broadening, strain, phonon-bottleneck effect, etc.

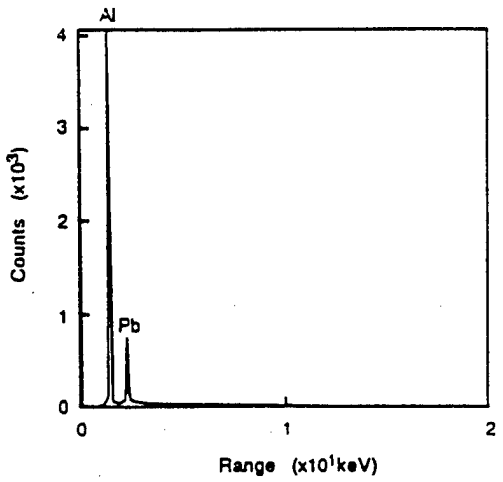



**Figure 15.** Energy dispersive x-ray data showing the presence of Pb in the pores. The samples were prepared by electrodepositing Pb into the pores from a solution of lead acetate.

### 4.2. Optical characterization

The photoluminescence data for CdS dots are shown in figure 16. Because of the relatively poor quantum efficiency and the dimness of photoluminescence, the spectrum is noisy. However, it shows definite indication of quantum confinement since the peak is blue shifted from the bulk CdS bandedge by 200 meV (bulk CdS bandgap at the measurement temperature of 14 K is ~ 2.41 eV) [18]. If we neglect any strain-induced shift or any Burstein–Moss shift, then this amount of blue shift indicates a quantum dot diameter of ~ 6 nm (in this calculation, we assumed that the electron and heavy-hole effective masses retain their bulk values; $m_e^* = 0.21m_0$ and $m_{hh}^* = 0.8m_0$). This 6 nm diameter is somewhat less than what quantitative image analysis indicates†. The latter indicates that the *physical* quantum dot diameter is $\sim 13 \pm 1$ nm. However, if the side depletion of the dots (caused by Fermi level pinning) is a mere 3.5 nm, then the *effective* optical dot diameter measured by this technique will be indeed 6 nm. A mere 3.5 nm side depletion is quite plausible for these systems.

A disturbing feature in the photoluminescence spectrum is its very broad width (400 meV full width at half maximum). There are at least four possible sources for this broadening.

- Homogeneous broadening caused by either impurities, surface states or by the fact that the CdS quantum dots are amorphous and not crystalline. The dots were annealed for six hours at 150 °C to recrystallize them. However, this may not have been sufficient. We are investigating this possibility using electron diffraction to ascertain the crystallinity (or lack thereof) of the CdS dots. However, the most likely culprit is *strain*. These dots are severely strained which may cause very large broadening.
- Inhomogenous broadening in the system. We found that even a 2 nm variation in the *effective* dot diameter will result in an inhomogeneous broadening of ~150 meV. This problem is pathological to extreme quantum confined structures [8]. The laser spot size used in exciting the dots has a diameter of 100 $\mu$m so that $4\times10^6$ dots are excited simultaneously. In future, we will etch mesas of area 10 $\mu$m$\times$10 $\mu$m and excite them individually to reduce the inhomogeneous broadening.
- Because of selection rules imposed by restricted momentum conservation in quantum dots, scattering of photoexcited carriers from higher subbands to the lowest subband is inhibited. This so-called 'phonon bottleneck' effect [19] can cause broadening of the spectrum since it promotes transitions involving excited states.
- Finally, it is possible that the broadening is caused by nonlinear optical effects. Since the amount of CdS is very small, it is necessary to use intense excitation so that the photoluminescence is detectable. If nonlinear effects are manifested at this excitation power (and they are expected to be), then harmonic generation may cause a wide spectrum. We are planning to study nonlinear effects by direct transmission/absorption

† If we account for any Burstein–Moss shift (which may not be negligible because the total number of states is small in a quantum dot and can be easily filled), then the quantum dot diameter will be larger than 6 nm.

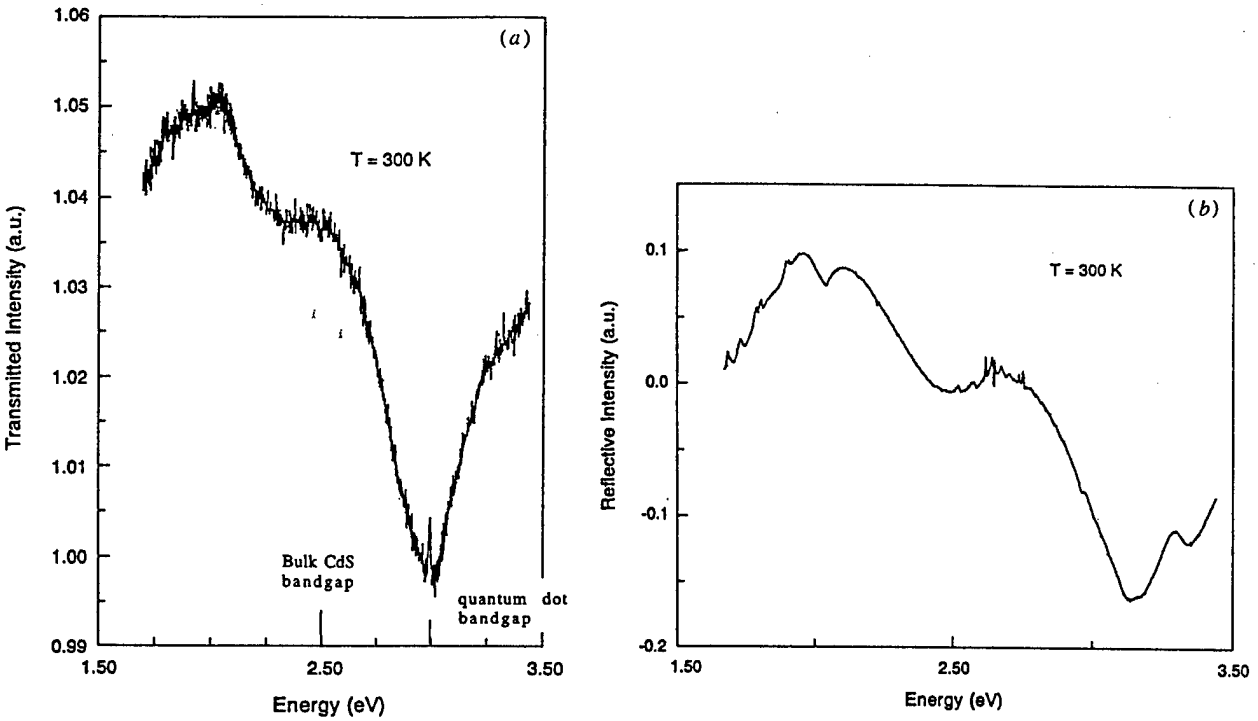



**Figure 17.** (*a*) Transmission and (*b*) reflectance spectrum showing a blue-shifted absorption edge at 3.0 eV which is 500 meV above the bulk bandedge. The calculated diameter of the quantum dots (assuming that the blue shift is entirely due to quantum confinement) is 3.5 nm.

measurements. These measurements will be carried out by stripping the Al substrate (after the CdS electrodeposition) using bromine/methanol solution and then capturing the stripped film on a transparent quartz slide. We expect to observe absorption saturation in these dots owing to either phase space filling, or bandgap renormalization, or inherent third order nonlinear susceptibility $\chi^{(3)}$ associated with formation of excitonic complexes (e.g. biexcitons). At the time of writing this manuscript, preliminary evidence of absorption saturation has been observed in pump–probe experiments by our collaborators.

Still smaller structures (produced by shorter duration of electrodeposition) have been characterized by absorption and reflectance studies. The spectra are shown in figure 17. In this case the blue-shift of the absorption bandedge is about 500 meV indicating an effective optical diameter of the dots to be 3.5 nm. These are some of the smallest semiconductor dots produced in non-agglomerated quasi-regimented arrays that are optically active.

In addition to photoluminescence, Raman spectroscopy of CdS quantum dots has also been performed. The data are shown in figure 18 where the longitudinal optical (LO) phonon peak for CdS and a peak at twice the LO frequency are consistently observed. The peaks were observed when the laser photon energy resonantly excited the Raman modes (2.71 eV or 2.54 eV) but were not observed when the photon energy was below the bandgap (2.41 eV). These data unambiguously indicate that the material formed within the pores is definitely CdS. Since the peaks were neither significantly wide nor asymmetric in their lineshape, we believe that the CdS material within the pores is mostly crystalline.

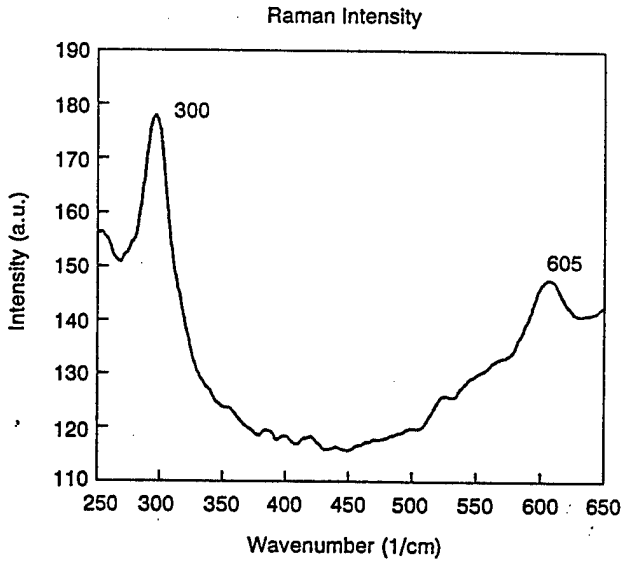



**Figure 18.** Raman spectra of CdS quantum dots showing the LO and 2LO phonon peaks. Other phonon peaks were not observed possibly because of inadequate quantum efficiency. Observation of these peaks characteristic of CdS indicates the presence of CdS within the pores.

### 4.3. Magnetic properties

To study magnetic properties, we fabricated an array of magnetic quantum dots using the structure shown in figure 14(*a*). This structure was then electroplated with copper and four leads were attached to the top copper layer. The four terminal resistance of this structure was measured in a magnetic field. Electrons flow in the copper layer but suffer spin-dependent scattering from the magnetic

moments of the underlying nickel quantum dots. Using a simple model, we predicted that this structure will exhibit giant magnetoresistance. Room temperature measurements indeed showed a 3% magnetoresistance up to a magnetic field of 2 T [14].

### 4.4. Superconducting properties

We have deposited the high-temperature superconductor YBCO within the pores from an organic electrolyte. Stoichiometry has been verified by Auger. Microwave surface resistance measurements have shown that these quantum dot superconductors retain the superconductivity and reach zero resistance at a slightly depressed $T_c$ of 88 K.

## 5. Conclusion

In conclusion, we have described two electrochemical self-assembly processes that we believe may lead to cheap and versatile techniques for producing regimented quantum dot arrays. Further characterization of the structures and theoretical simulations geared towards understanding the self-assembly mechanism are currently underway.

## Acknowledgments

This work was supported by the Department of Energy grant DE-FG02-90ER45427 administered by the Midwest Superconductivity Consortium (Nebraska, Notre Dame and Missouri), and by the US Army Research Office under grants DAAH04-95-1-0586 (Nebraska and Notre Dame) and DAAL03-92-G0381 and EPSCoR DAAL03-92-G0367 (Missouri). The work at Argonne was supported by the US Department of Energy under grant BES-DMS-W-31-109-ENG-38. The authors are grateful to Professor Howard Blackstead for microwave measurements of superconducting samples.

## References

[1] Das B, Subramaniam S and Melloch M R 1993 *Semiconduc. Sci. Technol.* **8** 1347
[2] Tsui D C, Gossard A C and Dolan G J 1983 *Appl. Phys. Lett.* **42** 180
[3] Fink T, Smith D D and Braddock W D 1990 *IEEE Trans. Electron Devices* **ED 37** 1422
[4] Stellwag T B, Melloch M R, Cooper J A Jr, Sheppard S T and Nolte D D 1992 *J. Appl. Phys.* **71** 4509
[5] Knoedler C M 1990 *J. Appl. Phys.* **68** 1129
[6] Chang T P 1996 *Proc. Int. Conf. Quantum Devices and Circuits (Alexandria, Egypt, June 1996)* to appear
[7] Sakaki H 1992 *Surface Science* **267** 623
Meurer B, Heitmann D and Ploog K 1992 *Phys. Rev. Lett.* **68** 1371
Wang P D, Sotomayer-Torres C M, Benisty H, Weisbuch C and Beaumont S P 1992 *Appl. Phys. Lett.* **61** 946
Schmidt A, Forchel A, Straka J, Gyuro I, Speier P and Zielenski E 1992 *J. Vac. Sci. Technol.* B **10** 2896
[8] Bawendi M G *et al* 1990 *Phys. Rev. Lett.* **65** 1623
Leonard D, Krishnamurty M, Reaves C M, DenBaars S P and Petroff P M 1993 *Appl. Phys. Lett.* **63** 3203
Moison J M, Houzay F, Barthe F, Leprince L, André E and Vatel O 1994 *Appl. Phys. Lett.* **64** 196
Ledentsov N N *et al* 1995 *Proc. Int. Conf. Phys. of Semiconductors* ed D J Lockwood (Singapore: World Scientific) pp 1855
Heitz R *et al* 1996 *Appl. Phys. Lett.* **68** 361
Grundmann M 1996 *Adv. Solid State Phys.* **35** 123
[9] Andres R P, Bielefeld J D, Henderson J I, Janes D B, Kolagunta V R, Kubiak C P Mahoney and Osifchin R G *Science* to appear
[10] Andres R P, Bein T, Dorogi M, Feng S, Henderson J I, Kubiak C P, Mahoney W, Osifchin R G and Reifenberger R G 1996 *Science* **272** 1323
[11] Andres R P 1994 *Conf. on Ordered Molecular and Nanoscale Electronics (Kona, Hawaii, June 1994)*
[12] Diggle J W, Downie T C and Goulding C W 1969 *Chem. Rev.* **69**, 365
Huber C A, Huber T E, Sad Mqio, Lubin J A, Manalis S and Prater C B 1994 *Science* **263** 800
Moskovits M 1993 US Patent # 5,202,290, issued April 13
[13] Penner R M and Martin C R **59**, 2625
Brumlik C J, Martin C R and Takuda K 1992 *Anal. Chem.* **64** 1201
Klein J D *et al* 1993 *Chem. Mater.* **5** 902
[14] Bandyopadhyay S, Miller A E and Chandrasekhar M 1995 *Proc. SPIE* **2397** 11
Miller A E, Yue D-F, Banerjee G, Bandyopadhyay S, Ricker R E, Jones S and Eastman J A 1994 *Quantum Confinement: Physics and Applications* ed M Cahay *et al* (Pennington, NJ: The Electrochemical Society Inc.) p 166
Yue D-F, Banerjee G, Miller A E and Bandyopadhyay S 1996 *Superlat. Microstruct.* **19** 191
Yue D-F 1995 *PhD thesis* University of Notre Dame and references therein
[15] Martin C R 1994 *Science* **266** 1961
[16] Kazuaki Sakoda 1995 *Phys. Rev.* B **52**, 8992 and references therein
Rosenberg A, Tonucci R J, Lin H B and Shirley E L 1996 *Phys. Rev.* B **54** R5195
[17] Beale R *et al* 1985 *J. Cryst. Growth* **73** 622
Searson P C 1993 *J. Appl. Phys.* **72** 253
Kang Y 1993 *J. Electrochem. Soc.* **140** 2258
[18] Sze S M 1981 *Physics of Semiconductor Devices* 2nd edn (New York: Wiley)
[19] Bockelmann U and Bastard G 1990 *Phys. Rev.* B **42** 8947
[20] Bandyopadhyay S and Roychowdhury V P 1995 *Phys. Low Dimensional Structures* **8/9** 29

*Superlattices and Microstructures, Vol. 22, No. 3, 1997*

# Switching in a reversible spin logic gate

S. BANDYOPADHYAY†
*Department of Electrical Engineering, University of Nebraska, Lincoln, Nebraska, 68588-0511, USA*

V. P. ROYCHOWDHURY
*School of Electrical Engineering, Purdue University, West Lafayette, Indiana 47907, USA*

*(Received 15 July 1996)*

Theoretical results for the adiabatic switching of a reversible quantum inverter–realized with two antiferromagnetically coupled single electrons in adjacent quantum dots—are presented. It is found that a large exchange interaction between the electrons favors faster switching but also makes the timing of the read cycle more critical. Additionally, there exists an optimal input signal energy to achieve complete switching. Only for this optimal signal energy does the inverter yield an unambiguous, logically definite state. An experimental strategy for realizing circuits based on such gates in self-assembled arrays of quantum dots is briefly discussed.

© 1997 Academic Press Limited

**Key words:** quantum dots, single electronics, quantum computing.

## 1. Introduction

Research in nanoelectronic classical Boolean logic circuits derived from single electron interactions in quantum dots has been a busy field for the last few years [1–9]. A number of ideas have appeared in the literature [1–9] that visualize building dissipative (non-reversible) logic circuits based on Coulomb or exchange interaction between single electrons in arrays of quantum dots. Some of these schemes (e.g. [2]), however, are not only flawed, but they also violate the basic tenets of circuit theory. The individual logic devices have no isolation between input and output so that the input bit cannot even uniquely determine the output bit! (for a discussion of this issue see [3–6, 10]).

In this paper, we explore a different type of gate. It is a quantum mechanical gate that is reversible and non-dissipative. It should be contrasted with 'parametron-type' constructs that dissipate less than $kT \ln 2$ energy per bit operation [11], but are otherwise not entirely non-dissipative. While the bits in a parametron are c-numbers, the bits in the quantum gate to be described are true qubits and the time evolution of the system is unitary. For the sake of simplicity, we consider the smallest quantum gate possible, namely an inverter. It is fashioned from two antiferromagnetically coupled single electrons in two closely spaced quantum dots as envisioned in [3–5]. The equilibrium steady-state behavior of such a system has been investigated by Molotkov and Nazin [7, 8]. Here, we will explore the dynamic behavior and the unitary time evolution of this system in a non-dissipative and globally phase-coherent environment.

† On leave from the University of Notre Dame.

0749–6036/97/070411 + 06 $25.00/0 sm970365 © 1997 Academic Press Limited

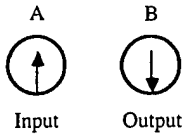



**Fig. 1.** Two adjacent quantum dots hosting single electrons. In the ground state, the spins of the two electrons are antiparallel. If spin polarization is used to encode binary bits, the logic state of one dot is always the inverse of the other. This realizes an inverter in which one dot acts as the input terminal and the other as the output.

## 2. Theory

Consider two single electrons housed within two closely-spaced quantum dots as shown in Fig. 1. It was shown in [3] that the preferred ordering of this system is antiferromagnetic, i.e. the two electrons have opposite spins. If the spin polarization in one dot is considered to be the input 'qubit' and that in the other the output 'qubit', then this system acts as an inverter since the spin-polarizations are antiparallel (logic complement) [3, 8]. Note that an inverter is always logically reversible since one can invariably predict the input bit from a knowledge of the output bit (in practice, the input bit is recovered by merely passing the output through another inverter). However, such a gate is not a universal quantum gate unlike the Toffoli gate [12]. Various schemes for realizing non-dissipative and reversible quantum logic gates have recently appeared in the literature [13–18]. Experimental demonstrations of quantum logic gates have also been reported [19, 20]. Almost all of these schemes encode the qubit in a photon (rather than an electron) state thereby requiring optical components that are incompatible with ultra-large-scale integration. In contrast, the spin gate based on single electrons in quantum dots is very appealing from the perspective of high-density circuits.

To analyse the system in Fig. 1 quantum-mechanically, we will assume that there is only one size-quantized level in each quantum dot. Then, the Hubbard Hamiltonian for this system in the presence of a globally applied magnetic field can be written following Molotkov and Nazin [7] as

$$\mathcal{H} = \sum_{i\sigma}(\epsilon_0 n_{i\sigma} + g\mu_B H_i \mathrm{sign}(\sigma)) + \sum_{\langle ij\rangle} t_{ij}(c^+_{i\sigma}c_{j\sigma} + h.c.) + \sum_i U_i n_{i\uparrow}n_{i\downarrow} + \sum_{\langle ij\rangle\alpha\beta} J_{ij}c^+_{i\alpha}c_{i\beta}c^+_{j\beta}c_{j\alpha} + H_z\sum_{i\sigma} g\mu_B n_{i\sigma}\mathrm{sign}(\sigma) \quad (1)$$

where the first term denotes the electron energy in the $i$th dot ($H_i$ is a $z$-directed local magnetic field selectively applied at the $i$th dot), the second term denotes the hopping between dots, the third term is the Coulomb repulsion within the $i$th quantum dot, the fourth term is the exchange interaction between nearest-neighbour dots and the last term is the Zeeman splitting energy corresponding to the globally applied magnetic field oriented along the $z$-direction.

We can simplify the Hamiltonian in Eqn (1) to the Heisenberg model following Molotkov and Nazin [8] to yield

$$\mathcal{H} = J\sum_{\langle ij\rangle}\sigma_{zi}\sigma_{zj} + J\sum_{\langle ij\rangle}(\sigma_{xi}\sigma_{xj} + \sigma_{yi}\sigma_{yj}) + \sum_{\text{input dots}}\sigma_{zi}h^{\text{input}}_{zi} \qquad (J > 0) \quad (2)$$

where we have neglected the global magnetic field. The quantity $h^{\text{input}}_{zi}$ is the Zeeman energy caused by a local magnetic field applied to the $i$th dot in the $z$-direction which will orient the spin in the $i$th dot along that field. Such a local field can be applied via a spin-polarized scanning tunneling microscope (SPSTM) tip as visualized in [3] and serves to provide an input signal to the gate.

**Table 1:** Eigenenergies and eigenstates of the Hamiltonian for an inverter.

| Eigenenergies | Eigenstates |
|---|---|
| $h_A + J$ | $\|\downarrow\downarrow\rangle$ |
| $-J + \sqrt{h_A^2 + 4J^2}$ | $\sqrt{\frac{1}{2}\left(1 + \frac{h_A}{\sqrt{h_A^2+4J^2}}\right)}\|\uparrow\downarrow\rangle + \sqrt{\frac{1}{2}\left(1 - \frac{h_A}{\sqrt{h_A^2+4J^2}}\right)}\|\downarrow\uparrow\rangle$ |
| $-J - \sqrt{h_A^2 + 4J^2}$ | $\sqrt{\frac{1}{2}\left(1 - \frac{h_A}{\sqrt{h_A^2+4J^2}}\right)}\|\uparrow\downarrow\rangle - \sqrt{\frac{1}{2}\left(1 + \frac{h_A}{\sqrt{h_A^2+4J^2}}\right)}\|\downarrow\rightarrow\rangle$ |
| $-h_A + J$ | $\|\uparrow\uparrow\rangle$ |

In the basis of states $|\sigma_A\sigma_B\rangle$ ($A$ and $B$ are the two electrons), the Hamiltonian in Eqn (2) can be written as

$$\begin{pmatrix} h_A + J & 0 & 0 & 0 \\ 0 & h_A - J & 2J & 0 \\ 0 & 2J & -h_A - J & 0 \\ 0 & 0 & 0 & -h_A + J \end{pmatrix} \tag{3}$$

where $h_A$ is the interaction with the input magnetic field selectively applied to quantum dot $A$. The two-electron basis states can be denoted as $|\downarrow\downarrow\rangle$, $|\uparrow\downarrow\rangle$, $|\downarrow\uparrow\rangle$ and $|\uparrow\uparrow\rangle$; they form a complete orthonormal set. The 'upspin' polarization is oriented along the direction of the locally applied external magnetic field in this representation.

The eigenenergies and corresponding eigenvectors of the above Hamiltonian are given in Table 1.

It is obvious that the third row in Table 1 corresponds to the ground state. In the absence of any external magnetic field ($h_A = 0$), the ground-state energy is $-3J$ and the ground-state wavefunction is $\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$. note that the ground state in the absence of any external magnetic field is an entangled state in which neither the quantum dot $A$ nor the quantum dot $B$ has a definite spin polarization.

## 3. Adiabatic switching

We now wish to study the following switching problem. Assuming that the inverter is in its ground state without any applied magnetic field, we will calculate how long it takes after a magnetic field is applied to quantum dot $A$ for the spin in $A$ to orient along the field and the spin in $B$ to orient in the opposite direction (as required by the inversion operation).

After the external field is applied at time $t = 0$, the inverter evolves in time according to the unitary operation

$$[c(t)] = \exp[-i\mathcal{H}t/\hbar][c(0)] \tag{4}$$

where $\mathcal{H}$ is given by Eqn (3) and $[c]$ is a four-element unit vector $[c_1, c_2, c_3, c_4]$ that describes the wavefunction $\psi(t)$ according to

$$\psi(t) = c_1(t)|\downarrow\downarrow\rangle + c_2(t)|\uparrow\downarrow\rangle + c_3(t)|\downarrow\uparrow\rangle + c_4(t)|\uparrow\uparrow\rangle. \tag{5}$$

The initial conditions are described by

$$\begin{bmatrix} c_1(0) \\ c_2(0) \\ c_3(0) \\ c_4(0) \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}. \tag{6}$$

The solution of Eqn (4) subject to the initial condition given by Eqn (6) is

$$c_1(t) = c_4(t) = 0$$
$$c_2(t) = \frac{e^{iJt/\hbar}}{\sqrt{2}}\left[\cos(\omega t) - i\left(\frac{h_A}{\hbar\omega} + \sqrt{1 - \frac{h_A^2}{\hbar^2\omega^2}}\right)\sin(\omega t)\right] \tag{7}$$
$$c_3(t) = -\frac{e^{iJt/\hbar}}{\sqrt{2}}\left[\cos(\omega t) - i\left(\frac{h_A}{\hbar\omega} - \sqrt{1 - \frac{h_A^2}{\hbar^2\omega^2}}\right)\sin(\omega t)\right]$$

where $\omega = \sqrt{h_A^2 + 4J^2}/\hbar$.

Therefore, the wavefunction at an arbitrary time $t$ is given by

$$c_2(t)|\uparrow\downarrow\rangle + c_3(t)|\downarrow\uparrow\rangle \tag{8}$$

with $c_2$ and $c_3$ given by Eqn (7).

After the switching is complete, the system should be in the state $|\uparrow\downarrow\rangle$. Therefore, the switching delay $t_d$ can be defined as the time taken for $|c_2(t)|$ to reach its maximum value and, correspondingly, for $|c_3(t)|$ to reach its minimum value.

This yields

$$t_d = \frac{h}{4\sqrt{h_A^2 + 4J^2}}. \tag{9}$$

It should be understood that the system *does not* reach a steady state at time $t = t_d$, but instead continues to evolve in accordance with Eqn (4). The computation (inversion) can be halted by reading the spin-polarization (logic bit) in the output dot (dot $B$) with a SPSTM tip at time $t = t_d$ since the reading operation is dissipative and collapses the wavefunction. Note that the higher the frequency $\omega$, the more critical is the timing for the read cycle that halts the quantum computation. Since $\omega$ increases with the exchange energy $J$, a larger $J$ will mandate a greater accuracy in the read cycle.

To achieve complete switching, the magnitude $|c_2(t_d)|$ should be unity and $|c_3(t_d)|$ should vanish. From Eqns (7) and (9), we obtain

$$|c_2(t_d)| = \frac{h_A + 2J}{\sqrt{2h_A^2 + 8J^2}}. \tag{10}$$

The magnitude $|c_2(t_d)|^2$ as a function of the normalized input signal energy $h_A/J$ is shown in Fig. 2. It reaches a maximum value of unity (corresponding to complete switching) when $h_A = 2J$. Therefore, there exists an optimal value of the input signal energy $h_A$ for which complete switching can be obtained.

It should be noted from Eqn (9) that the switching delay decreases with increasing exchange energy $J$. For the optimal case ($h_A = 2J$), the switching delay is $h/(8\sqrt{2}J)$. We can estimate the order of magnitude for $t_d$. Presumably, the maximum value of local magnetic field that can be applied to a dot with a SPSTM tip is about 1 T. Since $h_A \approx g\mu_B B$ ($\mu_B$ is the Bohr magnetron), this means that the maximum value of $h_A$ that we can hope to obtain is about 0.1 meV if we assume the Landé $g$-factor to be 2. Consequently, $J_{optimal} = 0.05$ meV. This gives a value of $t_d \approx 7$ ps. Therefore, these inverters are capable of quite fast switching.

We can also estimate the temperature of operation for such inverters. Since the exchange energy should exceed the thermal energy $kT$ for stable operation, the ambient temperature should be restricted to below $T = J/k = 570$ mK. Because the operation of the inverter requires global phase coherence (i.e. the phase breaking time should be significantly longer than $t_d$), a low temperature is also otherwise required. To increase the temperature to a more practical value of 4.2 K, $J_{optimal}$ should be 0.364 meV and therefore $h_A$ should be as large as 0.728 meV. This requires the ability to generate a local magnetic flux density in excess of 7 T with an SPSTM tip as an input to cell A. This is not possible with present state of SPSTM technology, but could become feasible in the future.

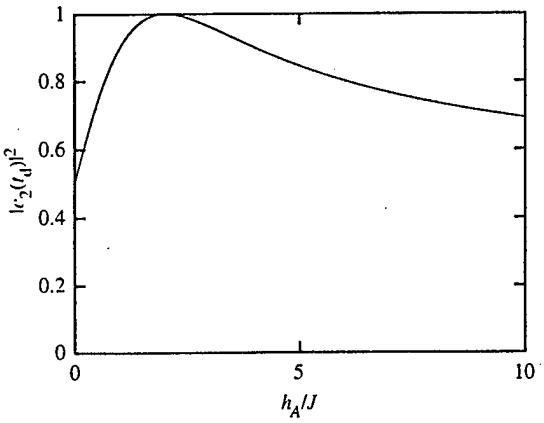



**Fig. 2.** The magnitude of $|c_2(t_d)|^2$ as a function of the normalized unput signal energy $h_A/J$.

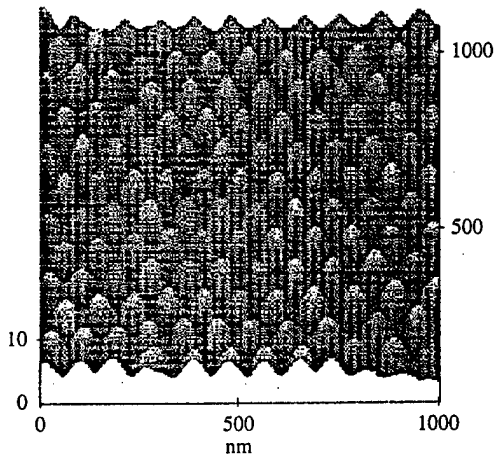



**Fig. 3.** Atomic force micrograph of a self-assembled mask to create a periodic array of quantum dots. Details can be found in [4, 6].

We conclude this paper with a brief discussion of experimental strategies undertaken by us in our efforts to fabricate such gates. We believe that the optimal technique is 'gentle' self-assembly of quantum dots rather than nanolithography which causes processing damage and has a slow throughput. We fabricate a regular array of the dots using a self-assembled mask for mesa-etching. The self-assembled mask is created by evaporating aluminum on the chosen semiconductor structure and then electropolishing it in a solution of perchloric acid, butyl cellusolve and ethanol at 60 V for 30 s at room temperature. Figure 3 shows the raw atomic force micrograph of a self-assembled mask of aluminum with a dimpled surface that consists of a periodic array of crests and troughs with hexagonal packing. The troughs are etched away by an appropriate etchant leaving a regular pattern of isolated crests on the surface of the semiconductor structure that serve as a mask through which mesas are etched. Owing to space limitations, we will omit details of the fabrication process, but instead refer the reader to [4, 6].

*Acknowledgements*—The work on self-assembly of masks for fabricating quantum dots was performed in collaboration with Prof. A. E. Miller, Prof. H. C. Chang, Dr. G. Banerjee and Mr. V. Yuzhakov of the Department

of Chemical Engineering, University of Notre Dame. This work was supported by the US Army Research Office under grant DAAH04-95-1-0586 and by the National Science Foundation under grant ECS9308814.

## References

[1] P. Bakshi, D. Broido, and K. Kempa, J. Appl. Phys. **70**, 5150 (1991).
[2] C. S. Lent, P. D. Tougaw, W. Porod, and G. H. Bernstein, Nanotechnology **4**, 49 (1993).
[3] S. Bandyopadhyay, B. Das, and A. E. Miller, Nanotechnology **5**, 113 (1994).
[4] S. Bandyopadhyay and V. P. Roychowdhury, Phys. Low Dimensional Structures **8/9**, 29 (1995).
[5] S. Bandyopadhyay and V. P. Roychowdhury, Jpn. J. Appl. Phys. part I **35**, 3350 (1996).
[6] V. P. Roychowdhury , D. B. Janes, S. Bandyopadhyay, and X. Wang, IEEE Trans. Elec. Dev. (Special Issue on Present and Future Trends in Device Science and Technologies, Vol. 43, 1688 (1996). Also Proc. of the IEEE, **85**, 574 (1997).
[7] S. N. Molotkov and S. S. Nazin, JETP Lett., to appear.
[8] S. N. Molotkov and S. S. Nazin, JETP, to appear.
[9] A. N. Korotkov, Appl. Phys. Lett. **67**, 2412 (1995).
[10] R. Landauer, Philos. Trans. R. Soc. London, **Ser. A** (in press).
[11] K. K. Likharev, Int. J. Theor. Phys. **21**, 311 (1982); K. K. Likharev and A. N. Korotkov, in *Proc. Int. Conference on Quantum Devices and Circuits* Alexandria, Egypt (1996) to be published.
[12] T. Toffoli, in *Automata, Languages and Programming*, eds. J. W. de Bakker and J. van Leeuwen Springer, New York, p. 632 (1980); Also see T. Toffoli, Int. J. Theor. Phys. **21**, 165 (1982); and E. Fredkin and T. Toffoli, *ibid*, **21**, 219 (1982).
[13] S. Loyd, Science **261**, 1569 (1993).
[14] A. Barenco, D. Deutsch, E. Ekert, and R. Josza, Phys. Rev. Lett. **74**, 4083 (1995).
[15] T. Sleator and H. Weinfurter, Phys. Rev. Lett. **74**, 4087 (1995).
[16] J. I. Cirac and P. Zoller, Phys. Rev. Lett. **74**, 4091 (1995).
[17] K. Obermeyer, W. G. Teich, and G. Mahler, Phys. Rev. **B37**, 8111 (1988).
[18] D. P. DiVincenzo, Phys. Rev. **A51**, 1015 (1995).
[19] Q. A. Turchette, C. J. Hood, W. Lange, H. Mabuchi, and H. J. Kimble, Phys. Rev. Lett. **75**, 4710 (1995).
[20] C. Monroe, D. M. Meekhof, B. E. King, W. M. Itano, and D. J. Wineland, Phys. Rev. **B75**, 4714 (1995).

*Superlattices and Microstructures, Vol. 19, No. 3, 1996*

# Giant magnetoresistance in an electrochemically synthesized regimented array of nickel quantum dots

D-F. YUE*, G. BANERJEE*, A. E. MILLER*, S. BANDYOPADHYAY†
**Department of Chemical Engineering, †Department of Electrical Engineering, University of Notre Dame, Notre Dame, Indiana 46556, USA*

*(Received 29 September 1995)*

Giant magnetoresistance (GMR) due to *remote* spin dependent scattering of electrons has been observed in an electrochemically synthesized structure consisting of a two-dimensional, quasi-periodic array of nickel dots (diameter ~100 Å) overlayed with a thin copper layer. Current flows exclusively in the copper layer, but the electrons scatter from the magnetic moments on the remote, underlying nickel quantum dots. Since the scattering cross-section depends on the magnetization of the dots, the resistance of the structure can be altered with a magnetic field which then gives rise to the GMR. The magnetoresistance is about 3% of the zero-field resistance up to a magnetic flux density of 2 tesla at room temperature.
© 1996 Academic Press Limited

## 1. Introduction

Giant magnetoresistance (GMR) is routinely observed in granular films consisting of a magnetic phase such as Co dispersed in a non-magnetic matrix such as Cu [1–3]. The GMR in these samples arises from the fact that the scattering cross-sections of electrons depends on the magnetization of the magnetic particles (Co). Since a magnetic field can alter the magnetization, the scattering rates and the resistance of the film can be modulated by a magnetic field which then gives rise to GMR.

Typically, granular films for GMR are prepared by dc sputtering from separate sources of immiscible ferromagnetic and non-magnetic materials [1]. This results in a very wide distribution in the magnetic particle size and also random placement of the particles within the film. Such variations make it difficult to study systematically the physics of the spin-dependent interaction of electrons with the magnetic moments. We therefore decided to prepare films that have a uniform distribution of ferromagnetic particles with well-controlled size. Accordingly, we prepared a structure consisting of a regimented quasi periodic array of Ni quantum dots with diameter $13 \pm 1$ nm, overlayed with a thin layer of Cu. Electrons mostly carry current through the Cu layer without encountering the Ni quantum dots. Nonetheless, they suffer remote scattering from the magnetic moments of the underlying Ni dots which then results in a GMR. To our knowledge, this is the first study of GMR due to *remote* scattering and also the first study of GMR in a film containing regimented arrays of size-controlled ferromagnetic particles.

## 2. Sample preparation

Samples are prepared in the following way. A foil of 99.999% pure aluminum is first electropolished in perchloric acid, ethanol and butyl cellusolve with 60 V for 30 seconds at 15°C. It is then anodized

0749–6036/96/030191+05 $18.00/0
© 1996 Academic Press Limited

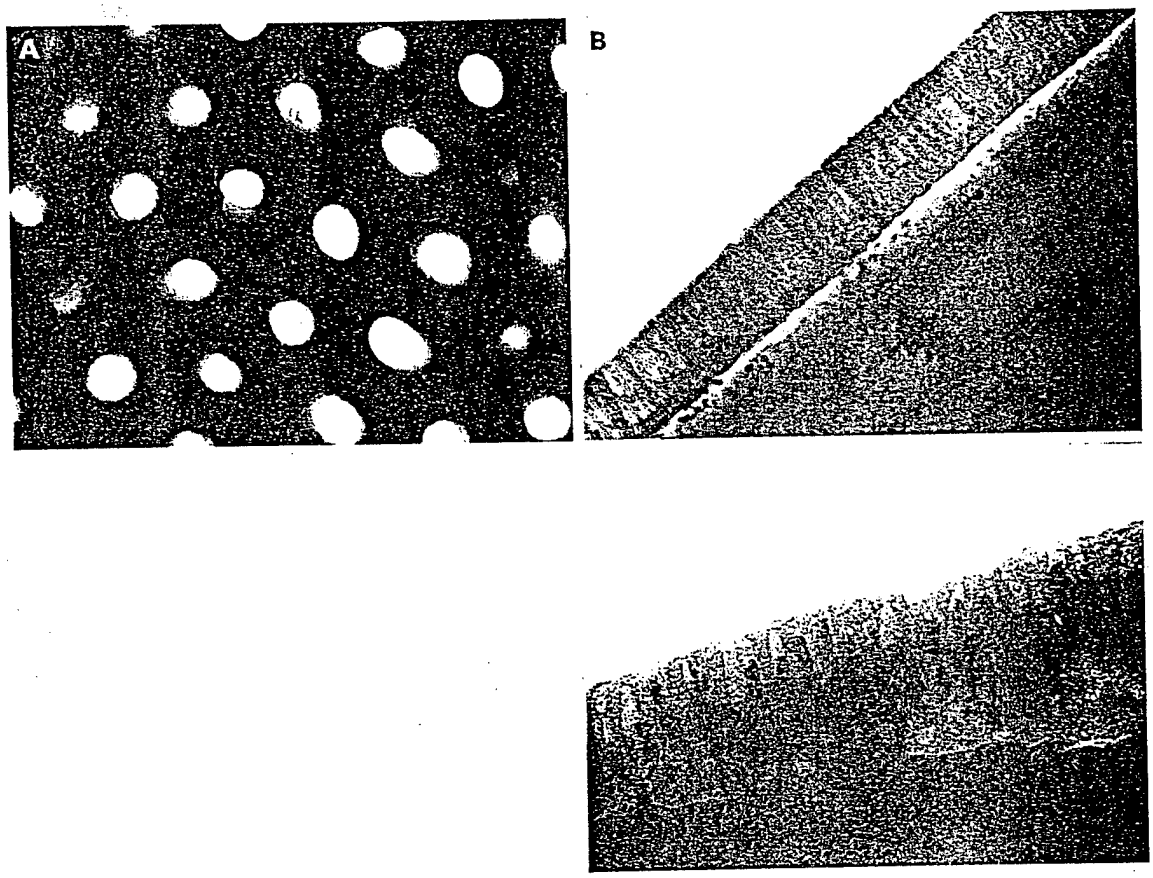



Fig. 1. A, Bright field TEM image of pores produced in anodized alumina film. The anodization was carried out with a dc current density of 40 mA $cm^{-2}$ in 15% $H_2SO_4$ (pH = 0.6) at room temperature. The pore diameter is $\sim 20 \pm 2$ nm. The pore density is $\sim 10^{11}$ $cm^{-2}$. B, Cross-section TEM of the porous film with electrodeposited Ni. The striations are the pores. In the top micrograph, the electrodeposition was carried out for 5 s resulting in quantum dots of diameter $\sim 13$ nm (the black dots in the pores). A 20 s deposition (bottom micrograph) results in the formation of vertical quantum wires with 180 nm height.

at room temperature in 15% $H_2SO_4$ (pH = 0.6) using a dc current density of 40 mA $cm^{-2}$. The anodization produces a nanoporous alumina film on the surface with a quasi-periodic distribution of pores [4] with $\sim 15$ nm diameter and a pore density of about $10^{11}$ $cm^{-2}$ [5]. The time of anodization determines the thickness of this porous film. A bright field TEM image of a typical alumina film is shown in Fig. 1A.

The pores are then filled up by Ni up to various depths using ac electrodeposition. For this, the porous film is washed in distilled water and dried. It is then placed in a non-cyanide acidic bath containing $NiSO_4$. AC current is passed at 18–20 V rms using the Al substrate and a platinum mesh as electrodes. The $Ni^{++}$ is preferentially electrodeposited in the pores since the pores offer the least resistance paths between the two electrodes [6]. The time of electrodeposition determines the depth of pore filling. One can monitor the surface resistance during electrodeposition, or carry out real time impedance spectroscopy, to ascertain the extent of pore filling. The surface resistance correlates very well with the filling and can be used to calibrate this process. For preparing GMR samples, we carried out the electrode position for four different durations of 3, 5, 10 and 15 seconds.

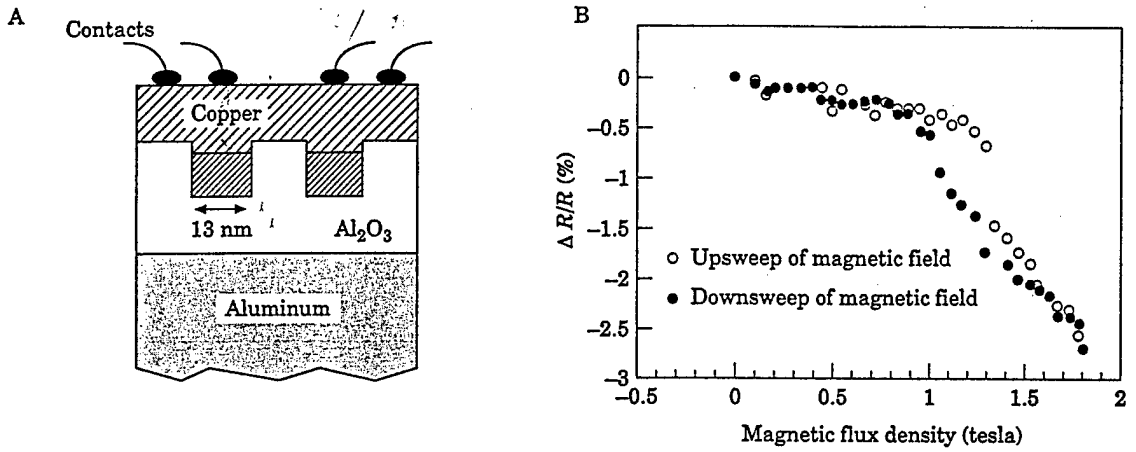



**Fig. 2.** A, A schematic cross-section of the sample used to measure GMR. B, The GMR data measured at room temperature.

## 3. Characterization of samples

To ascertain that we can indeed selectively electrodeposit Ni within the pores, we carried out cross-section TEM characterization. For this, we produced relatively thick porous alumina films (~200 nm) and electrodeposited Ni in the pores for two different durations of 5 and 20 s. We then produced specimens for cross-sectional TEM characterization using ultra-microtomy. TEM micrographs are shown in Fig. 1B. It is clear that a 5 s deposition produces fairly spherical dots of diameter 13 nm while a 20 s deposition produces short vertical wires of ~180 nm height. Samples that were used in the GMR experiments actually had a much thinner alumina film since the anodization was carried out for a much shorter duration. In that case, a 20 s Ni deposition would have certainly caused the Ni to overflow and spill outside the pores. Therefore, we restricted the Ni deposition to 15 s or less.

## 4. Experiment

Four-probe resistance measurements were made on four sets of samples with 3, 5, 10 and 15 s Ni deposition. Only the samples where the electrodeposition was carried out for 5 s showed significant GMR. The magnetoresistance data obtained at room temperature is shown in Fig. 2 along with the cross-section of the sample. Because of equipment limitations, the magnetic field was not scanned past 2 tesla and the magnetoresistance shows no sign of saturating at this field strength. Nonetheless, the value of the magnetoresistance is 3% of the zero-field resistance.

## 5. Analysis

Following Berkowitz *et al.* [1], the conductivity $G$ of this system measured between the two contacts can be written as

$$G=\frac{ne^2\tau}{m}, \tag{1}$$

where $n$ is the electron concentration in the Cu layer, $m$ is the electron's effective mass in Cu and $\tau$ is the momentum relaxation time or transport lifetime. Following Mathiesson's rule, we write

$$\frac{1}{\tau}=\frac{1}{\tau_{Cu}}+\frac{1}{\tau_{Ni}}, \tag{2}$$

where the first term is the characteristic momentum relaxation rate in Cu and the second term is the spin-dependent momentum relaxation rate associated with scattering from the isolated magnetic moments of the underlying ferromagnetic Ni dots. The second term is given by

$$\frac{1}{\tau_{Ni}}=\zeta(1\pm\lambda\vec{\sigma}\cdot\vec{M}_{Ni}), \tag{3}$$

where $\vec{\sigma}$ is the Pauli spin vector for the electrons, $\vec{M}_{Ni}$ is the magnetization of the Ni quantum dots (normalized with respect to the saturation magnetization) and $\lambda$ is a parameter that is indicative of the strength of remote scattering (note that the ferromagnetic dots do not lie in the path of the electrons unlike in conventional granular films and immiscible alloys [1]). When no magnetic field is present, the normalized magnetization of the Ni dots (at room temperature) is zero. It reaches a value of unity ($\vec{M}_{Ni}=1$) at the saturation field $B_s$. At this field strength, the electrons whose spins are aligned parallel and anti-parallel with the field have different conductivities given by

$$G^{\pm}=\frac{ne^2}{m}\left[\frac{1}{\tau_{Cu}}+\zeta(1\pm\lambda)\right]^{-1} \tag{4}$$

where the + superscripts refers to spins aligned along the field and the − superscripts to spins anti-aligned.

The conductivity at the saturation magnetic field is

$$G(B=B_s)=\frac{n^+G^+ + n^-G^-}{n}=\frac{G^+ + \frac{n^-}{n^+}G^-}{1+\frac{n^-}{n^+}}, \tag{5}$$

where $n^+$ and $n^-$ are the concentrations of electrons with the two different spin orientations. Their ratio depends on the Zeeman splitting $g\mu_B B$ ($\mu_B \equiv$ Bohr magnetron, $B \equiv$ flux density, $T \equiv$ absolute (temperature) and is given by $n^-/n^+ \approx exp[-g\mu_B B/kT]$. Therefore we get

$$G(B=B_s)=\frac{ne^2}{m}\left[\frac{1/\tau_{Cu}+\zeta-\xi\chi tanh[g\mu_B B_s/2kT]}{(1/\tau_{Cu}+\zeta)^2-\zeta^2\lambda^2}\right] \tag{6}$$

while the conductivity at zero magnetic field is

$$G(B=0)=\frac{ne^2}{m}\frac{1}{1/\tau_{Cu}+\zeta} \tag{7}$$

As long as the argument of the hyperbolic tangent is relatively small, the magnetoresistance will be negative (i.e. $G(B=B_s)>G(B=0)$). However, at very large fields, the magnetoresistance may change sign and become positive. If such a behavior is observed, it may be used to estimate $\lambda$.

## 6. Conclusion

We have shown the existence of GMR due to *remote* spin-dependent scattering in a structure consisting of a regimented array of size-controlled magnetic particles. Experiments are currently underway to measure the susceptibility of these particles and establish their magnetic properties.

*Acknowledgements*—This work was partially supported by the US Army Research Office under grant number DAAH04-95-1-0586. The fabrication work was also supported by the US Department of Energy through the Midwest Superconductivity. Consortium under the grant DE-FG02-90ER45427.

## References

[1] A. E. Berkowitz, Phys. Rev. Lett. 68, 3745 (1992).
[2] J. Q. Xiao, J. S. Jiang and C. L. Chien, Phys. Rev. Lett. 68, 3749 (1992).
[3] P. Xiong, Phys. Rev. Lett. 3220 (1992).
[4] See, for example, M. Moskovitz, US Patent No. 5 202 290 issued April 13, 1993.
[5] A. E. Miller, in: *Quantum Confinement: Physics and Applications*, eds. M. Cahay, *et al.* The Electrochemical Society, Inc., Pennington, New Jersey: p. 166 (1994).
[6] C. A. Huber, T. E. Huber, M. Sadoqi, J. A. Lubin, S. Manalis and C. B. Prater, Science 263, 800 (1994).

Abstract Submitted
for the MAR97 Meeting of
The American Physical Society

Sorting Category: 17.b

**Dielectric Enhancement of the Oscillator Strengths and Binding Energies of Magneto-Excitons and Magneto-Biexcitons in Quantum Wires** SUPRIYO BANDYOPADHYAY, ALEXANDER BALANDIN, *Department of Electrical Engineering, University of Nebraska, Lincoln, NE.*, ALEXEI SVIZHENKO, *Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN.*
— We will report a partly variational calculation of magneto- exciton and biexciton (excitonic molecule) binding energies in a rectangular semiconductor quantum wire taking into account dielectric confinement effects. The discontinuity of the dielectric constant at the boundaries between the wire and surrounding medium causes redistribution of electric field which leads to higher binding energies of excitons and biexcitons. The dielectric confinement effects are introduced via the image-charge method. Strong dielectric enhancement of the exciton binding energies (up to 2.5 times) and oscillator strengths is found for quantum wires of the dimensions comparable to the exciton Bohr radius. The enhancement decreases substantially in a magnetic field because of a charge cancellation effect. This work was supported by the US Army Research Office under contract DAAH04-95-1-0586 and DAAH04-95-1-0527.

Alexander Balandin
abalandi@engrs.unl.edu
Department of Electrical Engineering
University of Nebraska, Lincoln, NE.

[X] Prefer Oral Session
[ ] Prefer Poster Session

Date submitted: December 3, 1996

Electronic form version 1.2

Abstract Submitted
for the MAR97 Meeting of
The American Physical Society

Sorting Category: 17.b

**Exciton-Polariton Transport in an Array of Semiconductor Quantum Wires** ALEXANDER BALANDIN, SUPRIYO BANDYOPADHYAY, *Department of Electrical Engineering, University of Nebraska, Lincoln, NE.* — Polariton effects play an important role in low temperature absorption of bulk and nanostructured semiconductors in the spectral range close to the exciton resonance frequencies. The formation of polaritons (exciton-photon coupled states) modifies the transport of light through the medium. In particular, the medium of propagation becomes substantially more transparent provided that the exciton-polariton coherence length, critical temperature and exciton decay parameter meet certain criteria. We will show that in a properly designed array of quasi one-dimensional structures (quantum wires) of dimensions comparable to the exciton Bohr radius, the application of an external magnetic field can strongly increase the critical exciton-polariton coherence length and the critical temperature. Our numerical results are based on a partially variational calculation of the exciton longitudinal-transverse splitting as a function of wire dimensions and magnetic flux density. The temperature dependence is evaluated under the assumption that interactions with acoustic and polar optical phonons are the dominant scattering mechanisms for excitons. This work was supported by the US Army Research Office under contract DAAH04-95-1-0586 and DAAH04-95-1-0527.

Alexander Balandin
abalandi@engrs.unl.edu
Department of Electrical Engineering
University of Nebraska, Lincoln, NE

[X] Prefer Oral Session
[ ] Prefer Poster Session

Date submitted: December 3, 1996

Electronic form version 1.2

*Superlattices and Microstructures, Vol. 19, No. 2, 1996*

# Quasi one-dimensional excitons in a quantum wire subjected to a magnetic field

A. BALANDIN, S. BANDYOPADHYAY
*Department of Electrical Engineering. University of Notre Dame, Notre Dame, Indiana 46556, USA*

*(Received 29 September 1995)*

We have calculated the binding energy and 'length' of a quasi one-dimensional exciton in a semiconductor quantum wire subjected to a magnetic field. The magnetic field causes these quantities to have *non-monotonic* dependences on wire width. Below a critical width, the binding energy decreases with increasing wire width as usual, but above this critical width the opposite behavior is observed. This gives rise to a pronounced minimum in the binding energy. A similar behavior gives rise to a pronounced maximum in the exciton length. These features are explained by invoking the complementary roles of the electrostatic and magnetostatic confinement of the exciton.

© 1996 Academic Press Limited

## 1. Introduction

Quasi one-dimensional excitons and biexcitons in quantum wires are dirctly responsible for the giant third order non-linear susceptibility $\chi^{(3)}$ in these systems. The large magnitude of $\chi^{(3)}$ is caused by the increased binding energy of excitonic complexes due to one-dimensional confinement. A magnetic field can further enhance the confinement, leading possibly to improved performance in non-linear optics, specifically in low power and high density systems. Additionally, the field can act as an agent to modulate the non-linear absorption/gain in quantum wires which opens up the possibility of realizing *externally tunable* couplers, limiters, phase shifters, switches, etc.

Recently, Someya, Akiyama and Sakaki [1] reported the effect of an external magnetic field on the exciton binding energy and radius in a GaAs quantum wire by measuring the photoluminescence spectra and comparing them with those of quantum wells. They found that a magnetic field squeezes the exciton wavefunction to a size that is far below what can be achieved in quantum wells. This is consistent with our observation. We provide a detailed theoretical calculation which sheds light on this effect.

This paper is organized as follows. In Section 2, the theory of a quasi 1D exciton subjected to a magnetic field is developed rigorously within the framework of a two band model and perfect confinement. Section 3 presents the results of the variational calculations of the binding energy and exciton 'length' followed by a discussion of the excitonic properties. Conclusions are given in Section 4.

## 2. Theory

### 2.1 Exciton binding energy

Let us consider a quantum wire as shown in the inset of Fig. 1 with infinite potential barriers located at $y = \pm L_y/2$ and $z = \pm L_z/2$. A magnetic field of flux density $B$ is applied along the $z$-direction.

0749–6036/96/020097+08 $18.00/0 © 1996 Academic Press Limited

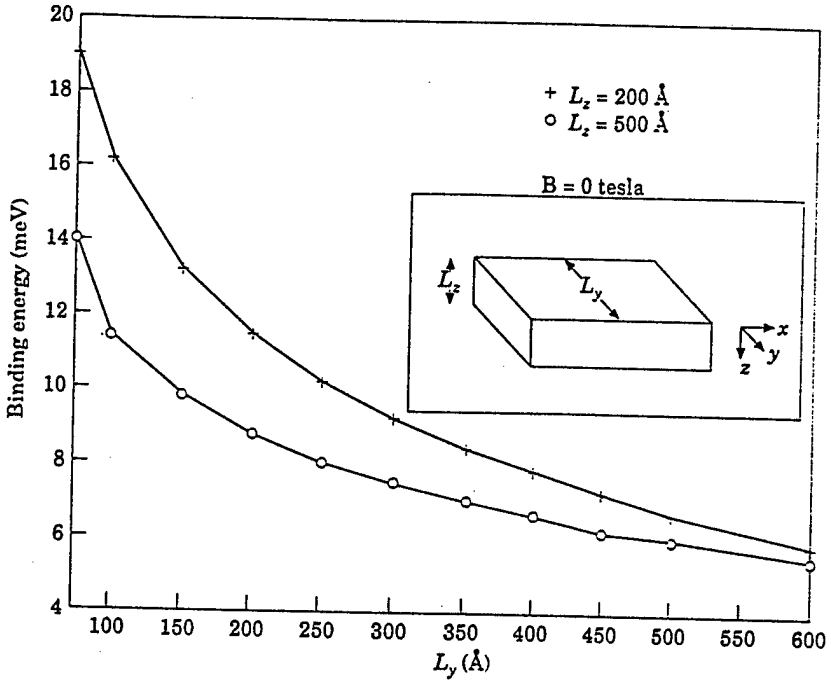



Fig. 1. Exciton binding energy in a GaAs quantum wire as a function of wire width $L_y$ for two different value of wire thickness $L_z$. No magnetic field is present. The binding energy decreases with increasing wire widt approaching the bulk value in sufficiently wide wires. The decrease is more rapid in the thicker wire because c the weaker confinement along the thickness. The inset shows a wire with the various coordinate direction defined.

For nondegenerate and isotropic bands, the Hamiltonian of a free Wannier exciton in thi system is given within the envelope-function approximation by

$$\hat{H}^X = \frac{1}{2m_e}(\hat{p}_{x_e} - eBy_e)^2 + \frac{1}{2m_h}(\hat{p}_{x_h} + eBy_h)^2 + \frac{\hat{p}_{y_e}^2 + \hat{p}_{z_e}^2}{2m_e} + \frac{\hat{p}_{y_h}^2 + \hat{p}_{z_h}^2}{2m_h}$$
$$- \frac{e^2}{4\pi\varepsilon[(x_e - x_h)^2 + (y_e - y_h)^2 + (z_e - z_h)^2]^{1/2}} + V_{conf}(y_e, y_h, z_e, z_h) \qquad (1$$

where we have chosen the Landau gauge for the magnetic vector potential:

$$\vec{A} = (-By, 0, 0).$$

The quantities $m_{e,h}$ and $x_{e,h}$, $y_{e,h}$, $z_{e,h}$ are the effective masses and coordinates of electrons and holes respectively, $\epsilon$ is the dielectric constant, $V_{conf}(y_e, y_h, z_e, z_h)$ is the confinement potentials for electrons and holes along $y$ and $z$ directions.

For convenience, we replace $x_{e,h}$-coordinates by the center-of-mass ($X$) and relative coordinates ($x$). This is accomplished by using quantum mechanical definition of momentum operators and taking into account that in a center-of-mass and relative coordinate system

$$\hat{p}_{x_{e,h}} = -i\hbar \frac{m_{e,h}}{M} \frac{\partial}{\partial X} \mp i\hbar \frac{\partial}{\partial x},$$

$$\hat{p}^2_{x_{e,h}} = -\hbar^2\left(\frac{m_{e,h}}{M}\right)^2\frac{\partial^2}{\partial X^2} \mp 2\hbar^2\frac{m_{e,h}}{M}\frac{\partial^2}{\partial X \partial x} - \hbar^2\frac{\partial^2}{\partial x^2}.$$

Defining

$$\hat{P}x \equiv -i\hbar\frac{\partial}{\partial X}$$

$$\hat{p}_x \equiv -i\hbar\frac{\partial}{\partial x}$$

we obtain

$$\hat{H}^x = \frac{\hat{P}x^2}{2M} + \frac{\hat{p}_x^2}{2\mu} + \frac{\hat{p}^2_{y_e}+\hat{p}^2_{z_e}}{2m_e} + \frac{\hat{p}^2_{y_h}+\hat{p}^2_{z_h}}{2m_h} + \frac{eB(y_e-y_h)}{M}\hat{P}x + eB(y_e/m_e + y_h/m_h)\hat{p}_x$$
$$+ \frac{e^2B^2}{2}(y_e^2/m_e + y_h^2/m_h) + V_{conf}(y_e,y_h,z_e,z_h) - \frac{e^2}{4\pi\varepsilon(x^2+(y_e-y_h)^2+(z_e-z_h)^2)^{1/2}}, \quad (2)$$

where

$$\frac{1}{\mu} = \frac{1}{m_e} + \frac{1}{m_h},$$

$$M\vec{R} = m_e\vec{r}_e + m_h\vec{r}_h,$$

$$\vec{r} = \vec{r}_e - \vec{r}_h,$$

$$\vec{R} = \hat{x}X + \hat{y}Y + \hat{z}Z,$$

Even for this relatively simple Hamiltonian of eqn (2), no exact analytical solution of the exciton wave function is possible. Therefore we adopt the standard variational approach [2–4]. Since the Hamiltonian does not depend on $X$, $P_X$ is a good quantum number. Dropping the term associated with $P_X$ we take the following trial wave function:

$$\psi \equiv \psi(x,y_e,y_h,z_e,z_h) = g_t(x,\eta)\phi_e(y_e)\phi_h(y_h)\chi_e(z_e)\chi_h(z_h), \quad (3)$$

where $g_t(x,\eta)$ is chosen to be the Gaussian-type 'orbital' function [6–8]:

$$g_t(x,\eta) = \frac{1}{\eta^{1/2}}\left(\frac{2}{\pi}\right)^{1/4} e^{-(x/\eta)^2} \quad (4)$$

in which $\eta$ is a variational parameter. The variables $X_{e,h}(Z_{e,h})$ are the $z$-components of the wave functions which are not affected by the magnetic field. They are given by particle-in-a-box states

$$\chi_{e,h}(z_{e,h}) = \sqrt{\frac{2}{L_z}}\cos\left(\pi\frac{z_{e,h}}{L_z}\right). \quad (5)$$

The electron and hole wave functions along the $y$ direction, $\phi_{e,h}(y_{e,h})$, are to be calculated numerically when a magnetic field is present. This is done by solving the Schrödinger equation directly following the prescription given in Ref. [11].

It is important to note that there are really two different cases of exciton quantization: (i) an electron-hole droplet whereby the exciton is considered to be a particle by itself, and (ii) an entity consisting of independently confined electron and hole. According to Ref. [9,10], the criterion for this separation is $L_y$, $L_z = 3a_B^*$, where $a_B^*$ is the effective Bohr radius in the bulk. The trial wave function (3) implicitly assumes the electron and the hole are independently confined along the $y$- and $z$-directions, which corresponds to the case

$$L_y, L_z < 3a_B^*.$$

The wave function (3) is probably the simplest that can be chosen while still preserving the principal features of the actual wave function. The results obtained with this trial function can be checked '*a posteriori*' by evaluating the zero-field binding energy and comparing it with binding energy calculated by other methods or extracted from experimental data.

The wave function (3) involves the variational parameter $\eta$ which is evaluated by minimizing the expectation value of the Hamiltonian in Equation (2) (based on given trial wave functions) with respect to $\eta$. Once this is accomplished, one can find the exciton binding energies and the exciton length for different values of magnetic field and the wire width. The functional to be minimized can be written as follows

$$\langle\psi|\hat{H}^X|\psi\rangle = \frac{\hbar^2}{2\mu\eta^2} + \frac{\hbar^2}{2\mu W^2} + \frac{\hbar^2}{2m_e}\int_{-L/2}^{L/2}(\phi_e')^2 dy_e + \frac{\hbar^2}{2m_h}\int_{-L/2}^{L/2}(\phi_h')^2 dy_h + \frac{e^2B^2}{2m_e}\int_{-L/2}^{L/2}(\phi_e y_e)^2 dy_e$$
$$+ \frac{e^2B^2}{2m_h}\int_{-L/2}^{L/2}(\phi_h y_h)^2 dy_h - \frac{e^2}{4\pi\varepsilon}\int_{\Omega}\frac{g_t^2(x,\eta)\phi_e^2\phi_h^2\chi_e^2\chi_h^2}{[x^2+(y_e-y_h)^2+(z_e-z_h)^2]^{1/2}}\,d\zeta, \tag{6}$$

where $d\zeta = dx dy_e dy_h dz_e dz_h$. The integration of the last (Coulomb) term is carried out over a hyperrectangle $\Omega$ which has an infinite interval along the $x$ direction and limited by $\pm L_y/2$ and $\pm L_z/2$ along $y$ and $z$ directions respectively. To obtain (6), we have made use of the boundary conditions $\phi_{e,h}(\pm L_y) = 0$ which allowed us to integrate some of the terms analytically using integration by parts. Note that the expectation value of the non-Hermitean operator $eB(y_e/m_e + y_h/m_h)\hat{p}_x$ (which arises in the presence of a magnetic field) is identically zero for the chosen trial wave function which makes the expectation value in (6) strictly real and shows that the trial wave function space is admissible.

Equation (6) allows us to treat the Coulomb interaction term exactly in its full 3D form throughout the calculation, which is physically more realistic than the approach employed in our previous (2D) exciton binding energy calculation [12].

Ground state exciton binding energies $E_B^X$ can now be found using the relation

$$E_B^X = E_{conf}^{e1} + E_{conf}^{hh1} - min\langle\psi|\hat{H}^X|\psi\rangle, \tag{7}$$

where $E_{conf}^{e1}$, $E_{conf}^{hh1}$ are the lowest electron and the highest heavy hole magneto-electric subband bottom energies in a quantum wire measured from the bottom of the bulk conduction band and the top of the bulk valence band. At the same time, the exciton length is simply $\eta_{opt}/2^{1/4}$ where $\eta_{opt}$ is the value of $\eta$ that minimizes the expectation value in eqn (6).

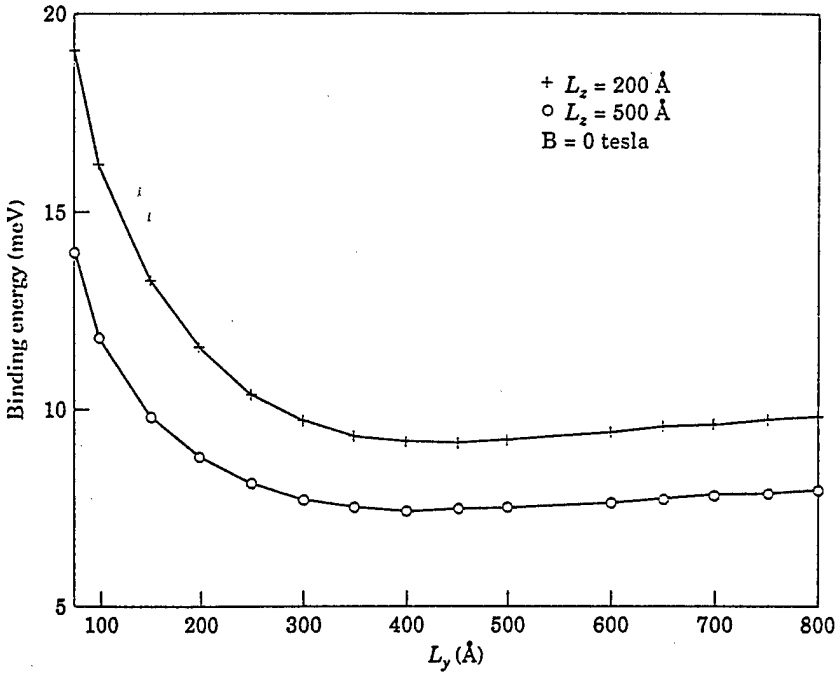



Fig. 2. The same results as in Fig. 2 in the present of a magnetic flux density of 10 tesla. There are pronounced minima around a wire width of 400 Å. The magnetic length in this case is 96 Å.

## 3. Results and Discussion

In Fig. 1 we present the exciton binding energy as a function of $L_y$ (wire dimension along $y$ direction) for two values of $L_z$ (wire dimension along $z$ direction) when no magnetic field is present. With increasing values of $L_y$, the binding energy rapidly decreases and begins to approach the bulk value (about 5.5 meV for GaAs). The binding energy does not approach the 2D limit because confinement along the $z$ direction is not strong enough. In Fig. 2 we display the same results when a magnetic flux density of 10 tesla is present. It is interesting to note that when a magnetic field is present, the binding energy curves have a clearly resolved *minima* at $L \approx 400$ Å. An explanation for this somewhat surprising behavior is provided later. Figure 3 show the exciton length as a function of wire width. Pronounced maxima at $L \approx 400$ Å are seen corresponding to the minima in the binding energy.

Unfortunately, we cannot examine the behavior of the binding energy or the exciton length in the limit $L \to \infty$ because our model does not contain any provision to make the transverse components of the wavefunction to deform into atomic Slater orbitals. However, a direct comparison of our results for zero magnetic field with those in Refs [6,13] shows excellent agreement.

Figure 4A and B shows the exciton binding energy as a function of the magnetic field for different values of $L_y$ and $L_z$. Binding energy increases with the magnetic field for all wire widths, which is in qualitative agreement with the results obtained for 2D systems [14,15], except that while the increase is sublinear in 2D systems, it is *superlinear* in 1D systems. This can be explained in two different ways. A magnetic field squeezes the electron and hole wavefunctions along all directions causing these states to condense into cyclotron (Landau) orbits whose radii shrink with increasing magnetic fields. As long as the wire width $W$ is comparable to the magnetic length or the lowest cyclotron radius $l$ ($=\sqrt{\hbar/eB}$), the additioal confinement induced by the magnetic field is not very important and the geometric (i.e. the electrostatic confinement of the walls) confinement predomi-

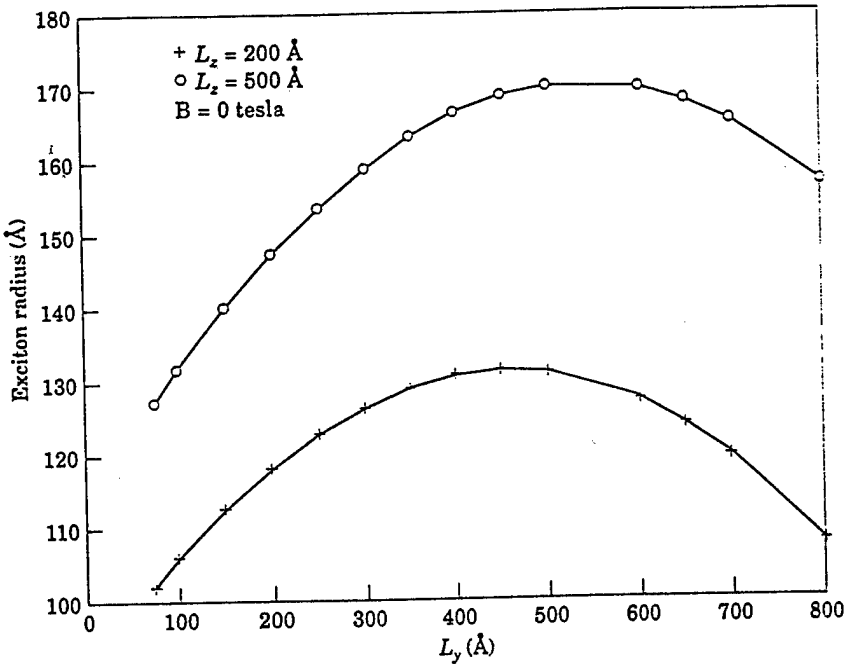



Fig. 3. Exciton length versus wire width at a magnetic flux density of 10 tesla. There are maxima around a wire width of 400 Å corresponding to the minima in the binding energy shown in Fig. 2.

nates. It is only when $W > l$ that the effect of the magnetic field becomes predominant. Therefore, a wider wire will show a stronger dependence of the binding energies on the magnetic field. Another way to explain the relation between the wire width and the sensitivity to the magnetic field is in terms of the standard time-independent perturbation picture. The magnetic field perturbs the quantum wire states, and the first order correction to the wave functions that correspond to the perturbed states is given by the formula

$$\psi_n^{(1)} = \sum_{m \neq n}^{\infty} \frac{|H'_{mn}|}{E_n^{(0)} - E_m^{(0)}} \psi_m^{(0)}, \tag{8}$$

where $E_m^{(0)}$, $\psi_m^{(0)}$ are the unperturbed energy eigenvalue and eigenfunction of the $m$th subband respectively, $|H'_{mn}|$ is a perturbation matrix element due to the magnetic field. Since in the case of perfect confinement

$$E_n^{(0)} = \frac{\pi^2 \hbar^2 n^2}{2m_{e,h} L_y^2}, \tag{9}$$

it is easy to see that the perturbation decreases with decreasing wire width. Since it is this perturbation that squeezes the exciton wave function, we see clearly that the squeezing is more effective in wider wires. In other words, the wave function is softer and more 'squeezable' in wider wires which causes the magnetic field effect to be more dominant in those wires. A very similar physics causes the hole wave function to be perturbed more than the electron wave function in a quantum wire [12].

We can now revisit the pronounced minima in Fig. 2 and try to explain their physical origin.

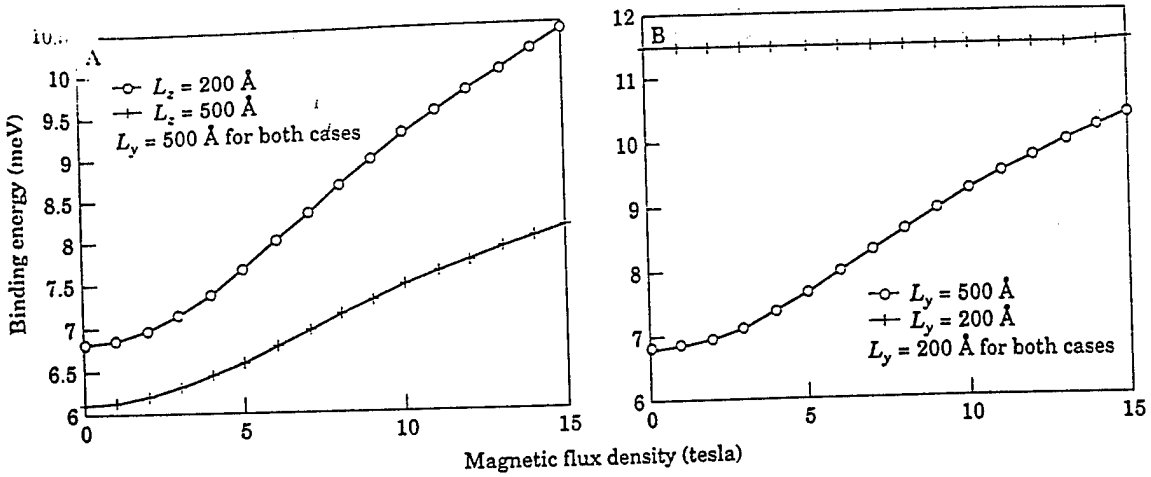



Fig. 4. Binding energy versus magnetic field for A, two different wire thicknesses and a fixed wire width, and B, two different wire widths and a fixed wire thickness.

nall values of $L_y$, the magnetic field is not very effective in squeezing the exciton wave function since the geometric confinement predominates. Consequently, the binding energy decreases with decreasing geometric confinement or increasing wire width. At large values of $L_y$, the geometric confinement becomes weak and yields the predominant role to the magnetic field induced confinement. With increasing wire width, the magneto-static confinement becomes stronger since the wave function becomes 'softer' and more squeezable. This causes the binding energy to reverse trend and increase with increasing wire width resulting in the occurrence of a minimum. The same physics explains the magnetic field dependence of the exciton radius as well.

## 4. Conclusion

In this paper, we have calculated the magnetic field dependence of the ground state exciton binding energy and exciton radius in a GaAs quantum wire. Two important observations are that: (i) the binding increases superlinearly with increasing magnetic field unlike in a quantum well where the increase is sub-linear [15], and (ii) there are pronounced mimima in the binding energy (and corresponding pronounced maxima in exciton length) as a function of wire width when a magnetic field is present. To our knowledge, the existence of these extrema was never shown before. Occurrence of these extrema has been explained in terms of the time-independent perturbation theory.

*Acknowledgements*—The authors are indebted to Dr Frank L. Madarasz for many illuminating discussions. This work was partially supported by the US Army Research Office under grant number DAAH04-95-1-0586.

## References

[1] T. Someya, H. Akiyama and H. Sakakai, Phys. Rev. Lett. 74, 3664 (1995).
[2] R. C. Miller, D. A. Kleinman, W. T. Tsang and A. C. Gossard, Phys. Rev. B 24, 1134 (1981).
[3] G. Bastard, E. E. Mendez, L. L. Chang and L. Esaki, Phys. Rev. B 26, 1974 (1982).
[4] D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Weigmann, T. H. Wood and C. A. Burrus, Phys. Rev. Lett. 53, 2173 (1984).

[5] D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Weigmann, T. H. Wood and C. A. Burrus, Phys. Rev. B **32**, 1043 (1985).
[6] F. L. Madarasz, F. Szmulowicz, F. K. Hopkins and D. L. Dorsey, J. Appl. Phys. **75**, 639 (1994).
[7] F. L. Madarasz, F. Szmulowicz, F. K. Hopkins and D. L. Dorsey, Phys. Rev. B **49**, 13 528 (1994).
[8] R. O. Klepfer, F. L. Madarasz and F. Szmulowicz, Phys. Rev. B **51**, 4633 (1995).
[9] A. D'Andrea and R. Del Sole, Phys. Rev. B **46**, 2363 (1992).
[10] P. Vicente, A. V. Kavokin, A. Raymond, S. G. Lyapin, K. Zekentes, D. Dur and W. Knap, Journal de Physique IV **3**, 323 (1993).
[11] S. Chaudhuri and S. Bandyopadhyay, J. Appl. Phys. **71**, 3027 (1992).
[12] A. Balandin and S. Bandyopadhyay, J. Appl. Phys. **77**, 5924 (1995).
[13] Marcos H. Degani and Oscar Hipolito, Phys. Rev. B **35**, 9345 (1987).
[14] R. Ferreira, B. Soucail, P. Voisin and G. Bastard, Superlattices and Microstructures **8**, 249 (1990).
[15] A. V. Kavokin, A. I. Nesvizhskii and R. P. Seisyan, Semiconductors **27**, 530 (1993).

# Giant dipole effect and second-harmonic generation in quantum wires biased with a magnetic field

A. Svizhenko
*Department of Electrical Engineering, University of Notre Dame, Notre Dame, Indiana 46556*

A. Balandin[a] and S. Bandyopadhyay[b]
*Department of Electrical Engineering, University of Nebraska, Lincoln, Nebraska 68588*

(Received 7 November 1996; accepted for publication 17 March 1997)

We have theoretically studied giant dipoles associated with transitions between magneto-electric subbands in a quantum wire subjected to a transverse magnetic field. The strengths of these dipoles and their resonant frequencies can be varied with the magnetic field which then allows one to tune the emission wavelength of these transitions. The large magnitude of the dipole moments also leads to a strong second-harmonic component of the dielectric susceptibility that can be utilized for nonlinear optical applications such as second-harmonic generation, limiting, mixing, optical switching, etc. © *1997 American Institute of Physics.* [S0021-8979(97)02612-1]

## I. INTRODUCTION

Direct intraband transitions between the quantized states (subbands) of the conduction band in a quantum well is a well-researched topic.[1] It has been shown both experimentally and theoretically that such transitions have very large dipole moments and narrow bandwidths. Strong infrared absorption, associated with transitions between the lowest two electronic subbands in a GaAs quantum well, was observed long ago by a number of experimental groups.[2] Recently, population inversion between the second and third subbands of a quantum well has been established unambiguously and has led to demonstration of the celebrated quantum cascade laser.[3] The energy separation between the subbands in a quantum well or wire can be varied by an external magnetic field which then allows one to realize a continuously tunable laser or light-emitting-device. Moreover, the field can induce forbidden transitions that make additional frequency ranges accessible, thus permitting flexible device design.

Another potential use of magnetic field biasing of quantum wells or wires is in nonlinear optics. Nonlinear optical properties stem from higher order dielectric susceptibilties. Specifically, the second-order susceptibility $\chi^{(2)}$ is responsible for such phenomena as mixing and second-harmonic generation. It is well known that even-order susceptibilities vanish in structures with inversion symmetry. Consequently, finite second-order susceptibilities can be obtained in such structures only if the inversion symmetry of the conduction-band potential is broken either by an external electric field or by the intentional growth of an asymmetric well. Obviously, the former is the preferred method since an electric field can be continuously varied which allows one to tune the degree of symmetry breaking and the magnitude of $\chi^{(2)}$. This method, however, has a practical shortcoming. An electric field tilts the potential barriers of the well thereby allowing carriers to escape by tunneling or thermionic emission. This is especially serious in GaAs/AlGaAs systems where the barrier height is relatively small. It has been pointed out that the electronic states in a quantum confined system biased by a transverse electric field are never true bound states since the particles can always lower their energy by escaping from the well.[4] Therefore, these states have a finite lifetime, which broadens the transitions.

To overcome this shortcoming, one can adopt magnetostatic biasing. A magnetic field can break inversion symmetry without tilting potential barriers and promoting carrier escape. A transverse magnetic field applied to a quantum wire exerts a Lorentz force on an electron moving along the wire. As a result, its wave function (in any subband) will be skewed towards one edge of the wire. This skewing does not tilt potential barriers to first order (the barriers may tilt slightly because of a second-order effect associated with space charges and the self-consistent (Hall) electric field). However, it effectively breaks inversion symmetry since it causes a net charge to accumulate at either edge of the wire (the charges at the two edges have opposite signs as in the classical Hall effect). This leads to a nonvanishing even-order susceptibility in a symmetric structure. The skewing has another subtle effect. The degree to which the wave function is skewed is *different in different subbands* since an electron has different kinetic energies and hence experiences different Lorentz forces in different subbands. As a result, transitions between subbands whose wave functions have the same parity — which are forbidden without a magnetic field — are now allowed since the parities are altered by different amounts in different subbands by the different degrees of skewing.[5]

This article is organized as follows. In Section II, we describe the theoretical formulation, followed by results. Finally, in Section IV, we present the conclusions.

## II. THEORY

We consider a quantum wire as shown in Fig. 1 with a magnetic field applied along the $z$ direction. The thickness

[a]On leave from the University of Notre Dame.
[b]Electronic mail: bandy@engrssl.unl.edu

 0021-8979/97/81(12)/7927/7/$10.00 © 1997 American Institute of Physics 

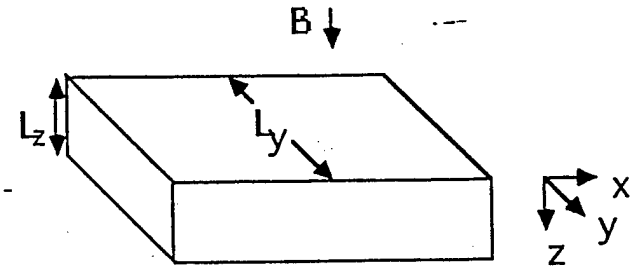



FIG. 1. An electron waveguide (quantum wire) subjected to a magnetic field along the $z$ axis. The width of the wire is much larger than the thickness.

along the $z$ direction is so small (and consequently the subband separation in energy is so large) that, for the range of photon energies considered, an electron cannot be excited (by real transition) into a subband which has more than two nodes along the $z$ direction. Such a transition will not be accessible in energy. This restriction, coupled with the fact that a magnetic field does not affect the $z$ component of the electron wave function, allows us to drop the $z$ component from further consideration. The width of the wire along the $y$ direction is however large enough that subbands with more than two nodes along the $y$ direction are accessible in energy.

In the framework of the envelope function approximation (EFA), an electron wave function can be written as the product of a Bloch wave function, periodic with the atomic lattice spacing, and an envelope wave function, describing the nonperiodic behavior. Consequently, the wave function of an electron for a given wave vector $k$ along the $x$ direction, in the $n$th magnetoelectric subband, at a magnetic field $B$ can be written as

$$\Phi_n(x,y,k,B,t)=\Psi_n(x,y,k,B)u_n(x,y,z,k)e^{-jE_n(k,B)t/\hbar}, \tag{1}$$

where $\Psi_n(x,y,k,B)$ is an envelope function, $u_n(x,y,z,k)$ is a Bloch function of a conduction band and $E_n(k,B)$ is the dispersion relation of the $n$th magnetoelectric subband at a flux density $B$. The Bloch wave functions are assumed to be $s$ states which is the usual case for semiconductors where $\hat{J}=1/2$ for the conduction band.

The envelope function can be further decomposed into a plane wave along the unconfined $x$ direction and a confined component along the $y$ direction. Thus,

$$\Psi_n(x,y,k,B)=\chi_n(y,k,B)e^{jkx}. \tag{2}$$

Using the electric dipole approximation, we can write the matrix element of photoinduced intersubband transitions within the conduction band as[6]

$$d_{f,i}(k,B)=e\int \chi_f(y,k,B)\,\hat{\eta}\cdot\vec{r}\,\chi_i(y,k,B)d\vec{r} \times \int u_f^*(x,y,z,k)u_i(x,y,z,k)d\Omega, \tag{3}$$

where $d\Omega$ is a volume element, $\hat{\eta}$ is the unit vector along the direction of the incident photon polarization, $\vec{r}=x\vec{a}_x+y\vec{a}_y$ is the two-dimensional radius vector, and subscripts $i,f$ stand for initial and final states respectively. The exponential term of Eq. (2) is not present in Eq. (3) since, for photoinduced transitions ($k_f=k_i$), the product of the exponential function and its complex conjugate are exactly unity. The volume overlap of the Bloch functions is also unity for $s$ states with the same wave vector. Now, if we assume that the incident light is polarized along the $y$ direction so that $\hat{\eta}=\hat{a}_y$, the above equation simplifies to

$$d_{f,i}(k,B)=e\langle\chi_f|y|\chi_i\rangle=e\int_{-W/2}^{W/2} y\chi_f(y,k,B)\chi_i(y,k,B)dy, \tag{4}$$

where $W$ is the width of the quantum wire along the $y$ direction.

One should note here that if there is no magnetic (or electric) field applied, the envelope functions $\chi_i$ are just particle-in-box states and the dipole moment in Eq. (4) is nonzero only for the transitions between subband states of opposite parity. For a symmetric square potential well, these dipole elements (between any two states $n$ and $m$) are independent of the wave vector $k$ and can be found analytically[1] by evaluating the integral in Eq. (4).

$$d_{f,i}=e\langle\chi_f|y|\chi_i\rangle=eW\frac{8}{\pi^2}\frac{mn}{(m^2-n^2)^2},$$

if $n$ and $m$ have opposite parity

$=0$, otherwise. (5)

However, when a magnetic field is applied, the skewing of the wave functions changes the integral in Eq. (4) and alters the selection rules. Generally, the skewing causes three effects. First, it makes the dipole moment depend on the wave vector $k$ (since the degree of skewing depends on $k$). Second, it reduces the dipole moment for transitions between states of opposite parity [since the integral in Eq. (4) decreases], and third, it allows forbidden transitions between states of the same parity [since the integral in Eq. (4) no longer vanishes for states of the same parity].

It is clear from Eq. (4) that, to calculate the dipole moments in the presence of a magnetic field, all we need to compute are the wave functions $\chi_{f,i}(y,k,B)$ at a given magnetic field $B$, for given magnetoelectric subbands $f$ and $i$, and for a given wave vector $k$. This is achieved via a numerical (finite difference) solution of the Schrödinger equation following the prescription of Ref. 7. Once this is done, we can calculate the dipole moment in Eq. (4) for any chosen intersubband transition at any chosen magnetic field and for any chosen wave vector.

In the limit of high magnetic fields, when the magnetic length $l(=\sqrt{\hbar/eB})\ll W$, one can again obtain an analytical expression for the dipole moment $d_{f,i}$. In this case, the magnetostatic confinement predominates over electrostatic confinement and the envelope functions $\chi_n(y,k,B)$ can be approximated by harmonic-oscillator wave functions:

$$\chi_n(y,k,B)\equiv\chi(y-y_k,B) =N_nH_n(\alpha,y-y_k)e^{-\frac{1}{2}\alpha^2(y-y_k)^2}, \tag{6}$$

where $N_n=(\alpha/\pi^{1/2}2^2n!)^{1/2}$ is a normalization constant, $H_n(\alpha,y)$ is the $n$th Hermite polynomial, $y_k=\hbar k/eB$, and

$$\alpha=\sqrt{\frac{eB}{\hbar}}=\frac{1}{l}. \tag{7}$$

In order to evaluate the integral in Eq. (4) analytically, we extend the limits of integration to infinity assuming that the wave function tail is negligible at the boundaries of the wire (i.e. at $y=\pm W/2$). This is a very reasonable assumption in a high confining magnetic field. The resulting analytical expression for the dipole moment is

$$d_{f,i}(B)=e\langle\chi_f|y|\chi_i\rangle=el\left(\frac{n+1}{2}\right)^{1/2}, \quad \text{if } m=n+1 \tag{8}$$

$$=el\left(\frac{n}{2}\right)^{1/2}, \quad \text{if } m=n-1$$

$$=0, \quad \text{otherwise.}$$

The physical significance of the two analytical limits, $B\to 0$ and $B\to\infty$, is obvious. At zero field, the dipole is determined by the width of the wire $d_{f,i}\sim eW$, and at the high field limit it is determined by the magnetic length $d_{f,i}\sim el$. This is what one would expect intuitively. At zero field, the dipole is confined electrostatically with the wire width being a measure of this confinement while at high magnetic field, the dipole is confined magnetostatically and the magnetic length is the corresponding measure of this confinement.

## III. RESULTS

### A. Intraband dipoles

We now present results of our calculations. The physical parameters used for the numerical calculations correspond to a GaAs quantum wire with relative dielectric constant $\epsilon_r=12.9$, and effective mass $m_e=0.067m_0$ where $m_0$ is the free electron mass.

In Fig. 2, we show the dependence of the dipole moment $d_{f,i}(k,B)$ for three transitions ($e1$-$e2$, $e2$-$e3$, and $e1$-$e3$) on the wave vector $k$ when a magnetic field of 1 T is applied (following usual practice, the transitions are numbered by the subband indices). The dipoles corresponding to transitions between states of opposite parity ($e1$-$e2$ and $e2$-$e3$) have maxima at $k=0$ and then decrease with increasing wave vector. This can be easily understood as follows. At zero wave vector (no translational velocity) these states do not experience any Lorentz force and hence the wave functions are not skewed. As the wave vector $k$ increases, the translational velocity and the Lorentz force experienced increase. Consequently, the envelope wave functions are skewed more and more and the dipole moment decreases. Real transitions between states of the same parity are forbidden at zero magnetic field, but at a finite magnetic field, they are forbidden only at $k=0$ when there is no translational velocity and no Lorentz force to skew the wave functions. With increasing $k$, the wave functions are increasingly skewed and the dipole moment of forbidden transitions increases. In our chosen

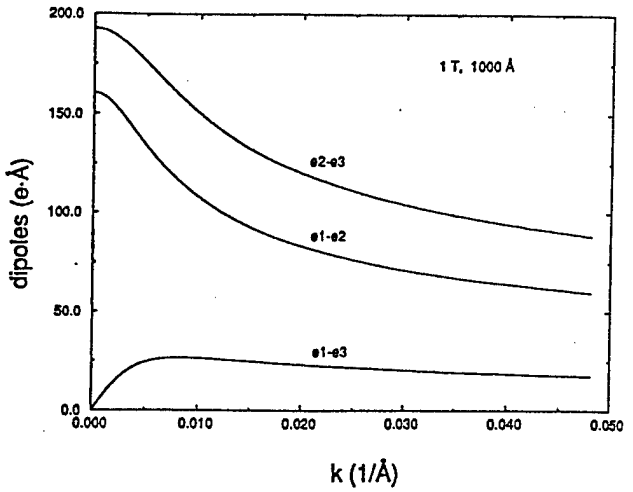



FIG. 2. Dipole moments for various intersubband transitions as functions of the propagating wave vector $k$ for a magnetic flux density of 1 T. At zero translational velocity ($k=0$) the dipole of transition $e1$-$e3$ vanishes. The GaAs quantum wire is 1000 Å wide.

prototype wire, $d_{e1-e3}$ reaches a maximum of 28 $e$-Å at $k=0.0051$ Å and then decreases gradually ultimately reaching zero. This intriguing *nonmonotonic* dependence on $k$ is explained later on. However, at this point, it is interesting to note that a fairly large forbidden dipole moment of ~30 $e$-Å can be achieved in realistic structures at a moderate magnetic field of 1 T.

Fig. 3 presents the dipole moments for the same transitions as a function of magnetic flux density. The propagation wave vector $k$ is chosen to be 0.01/Å. At zero magnetic field, a nonvanishing dipole matrix element occurs only for transitions between states of opposite parity ($e1$-$e2$, $e2$-$e3$) as expected from Eq. (5). This equation also allows us to estimate the strengths of these zero-field dipoles to be 180 $e$-Å for $e1$-$e2$ and 195 $e$-Å for $e2$-$e3$ transitions. As we can see from Fig. 3, these values are in excellent agreement with our

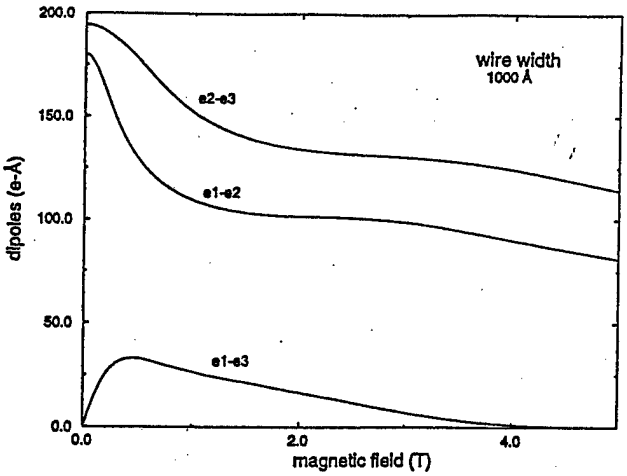



FIG. 3. The dipoles of three intersubband transitions as functions of the applied magnetic field. The dipole $d_{e1-e3}$ peaks at a magnetic flux density of 0.3 T. The wire width is the same as that in Fig. 2.

numerical results. From the analytical expression in Eq. (8), we can estimate the strength of the $e1$-$e2$ dipole to be 66 $e$-Å at a magnetic flux density of 15 T. This number also agrees with our numerical result. The $e1$-$e3$ dipole vanishes at both zero field (because of the spatial symmetry of the particle-in-a-box states) and at high fields because of the symmetry of the Landau states or Hermite polynomials. This behavior is consistent with Eqs. (5) and (8). Only at intermediate fields, when the wave functions of the subbands are a hybrid between particle-in-a-box states and Hermite polynomials (and thus "nonsymmetric" in space), is this transition allowed. This immediately tells us that $d_{e3-e1}$ must have a *nonmonotonic* dependence on the magnetic flux density $B$ and indeed it does.

Let us now examine the nonmonotonic behavior of $d_{e3-e1}$ more closely. This transition is forbidden at zero field since the wave functions of the first and third subbands have the same parity. At low magnetic fields, the parities are altered by the skewing of the wave functions and consequently $d_{e3-e1}$ is no longer zero but increases with the magnetic field. It reaches a maximum of about 30 $e$-Å and then decreases. This latter decrease is related to the following effect. For a fixed wave vector $k$, a sufficient increase in the flux density $B$ forces the traversing states ("skipping orbits" or "edge states") to condense into closed cyclotron orbits (Landau levels) that are no longer skewed by the magnetic field to the wire edge since they have no translational velocity and hence no Lorentz force. While edge states have a skewed wave function that is not symmetric in space, cyclotron orbits have a wave function that is symmetric about the orbit center $y_k$. Note that $y_k$ depends only on $k$ and $B$. Therefore, at a fixed $k$, the wave functions of the first and third Landau levels are symmetric about a *common* center. Whenever this kind of symmetry holds, $d_{e3-e1}$ vanishes. Therefore, the dipole moment $d_{e3-e1}$ decreases gradually to zero at high magnetic field with the onset of Landau condensation.

The same physics can be elucidated in a different way by considering the energy versus wave vector relation in Figs. 4(a) and 4(b) which show the dispersion of the first and third magneto-electric subbands respectively.

At $B=0$, velocity (slope of the curves) at $k=0.01$/Å are nonzero for both the $e1$ and $e3$ subbands. However, the Lorentz force is zero because $B=0$ and hence $d_{e1-e3}=0$. At $B=5$ T, the group velocities for the two subbands are still nonzero and the Lorentz force is finite resulting in skewing of wave functions and a nonvanishing value of $d_{e1-e3}$. At $B=10$ T, the group velocities at $k=0.01$/Å are zero in both subbands indicating that the corresponding states have undergone Landau condensation. In this case, the Lorentz force (for skewing) is again zero and the dipole moment $d_{e1-e3}$ vanishes once more. The crucial point to note is that the Lorentz force $e\vec{v}\times B$ can vanish in two different ways: (i) $B=0$, and (ii) $\vec{v}=0$. These two conditions are met at zero and very high magnetic fields. As a result, the dipole moment $d_{e1-e3}$ exhibits a nonmonotonic behavior in magnetic field. One can ask why the same physics does not cause nonmonotonicity in the $e1$-$e2$ and $e2$-$e3$ curves. It is not clear *a priori* that nonmonotonicity cannot occur (indeed

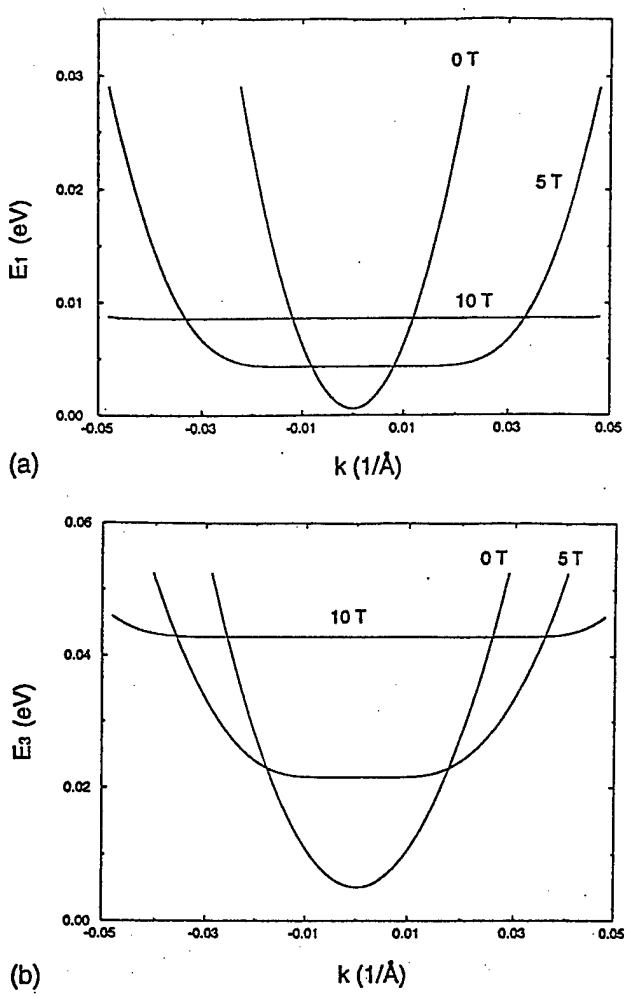



FIG. 4. Energy vs wave vector relation of electrons in (a) the first subband and (b) in the third subband of a 1000 Å wide quantum wire. The wave vector is along the free propagation direction. The results are shown for three values of a magnetic field. The energy is calculated from the bulk conduction band edge and the confinement energy for the $z$-direction is assumed to be zero.

there are regions of inflexion in the two curves). However, the point to note is that Landau condensation causes recovery of the wave function symmetry (or antisymmetry), *but does not restore the original zero-field wave functions*. This is shown in Fig. 5 where we show the wave functions in the $e1$ subband at 0 and 10 T. Both wave functions are "symmetric" in space, but they are otherwise vastly different since the magnetostatic confinement squeezes the wave functions binding them in cyclotron orbits.

The nonmonotonicity in the wave vector dependence of $d_{e1-e3}$ in Fig. 2 has a similar origin. As $k$ is increased, the *relative* skewing between the wave functions in $e1$ and $e3$ subbands change nonmonotonically causing the nonmonotonicity seen.

The process described above is illustrated in Figs. 6(a)–6(c), where we present wave functions of two electronic states ($e1$ and $e3$) for three values of magnetic flux density. At zero magnetic field the wave functions are symmetric about the center of the wire and dipole transition $d_{e3-e1}$ is forbidden [Fig. 6(a)]. At low magnetic field the wave functions are skewed to the edge of the wire ["edge states" Fig.

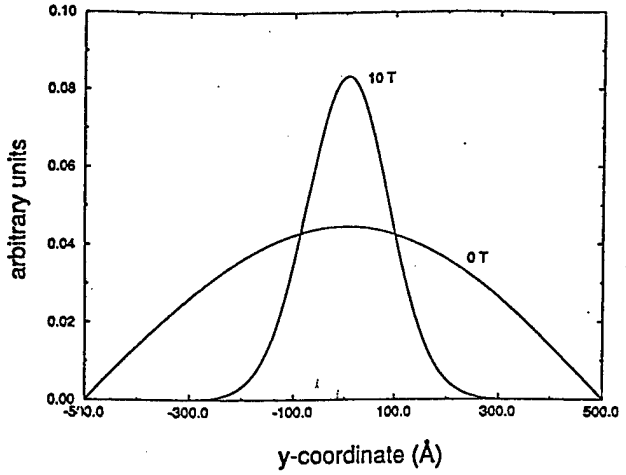



FIG. 5. The $y$ component of the electron envelope function for the first subband at a magnetic flux densities of 0 and 10 T.

6(b)] and the spatial symmetry is broken for both states. Dipole transition $d_{e3-e1}$ is now allowed. It is important to note here that the symmetry breaking skewing of the wave functions is caused by the simultaneous presence of a magnetic field and the electrostatic potential barriers at the edges of the quantum wire. At higher magnetic fields, when the magnetic length is smaller than the wire width, the electrons do not "feel" the potential barriers at the edges of the wire as they undergo complete Landau condensation and execute cyclotron motion with a radius much smaller than the width of the wire. In this case, the wave function symmetry is essentially restored [Fig. 6(c)] although the wave functions are now symmetric about a point that is not at the center of the wire. Nonetheless, what is important is that both wave functions are symmetric about the same point. Consequently, the $d_{e3-e1}$ transition vanishes. The simultaneous presence of both electrostatic confinement and magnetostatic confinement is therefore necessary for wave function skewing, formation of edge states and the observation of forbidden transitions.

## B. Second-harmonic generation

It is well known that in systems with inversion symmetry there can be no second order nonlinearity.[8] However, in systems without inversion symmetry, the lowest order optical nonlinearity is of the second order and is expressed by

$$\vec{P}^{(2)}(\vec{k},\omega)=\chi^{(2)}(\omega;\omega_1,\omega_2)\vec{E}_1(\vec{k_1},\omega_1)\vec{E}_2(\vec{k_2},\omega_2), \quad (9)$$

where $\vec{P}$ is the polarization caused by two electric fields $\vec{E}_1$ and $\vec{E}_2$ that are associated with the electromagnetic fields of either two frequency components of the same light beam or two different coherent beams with frequencies $\omega_i$ and wave vectors $\vec{k_i}$. The frequencies and wave vectors obey the energy and momentum conservation laws

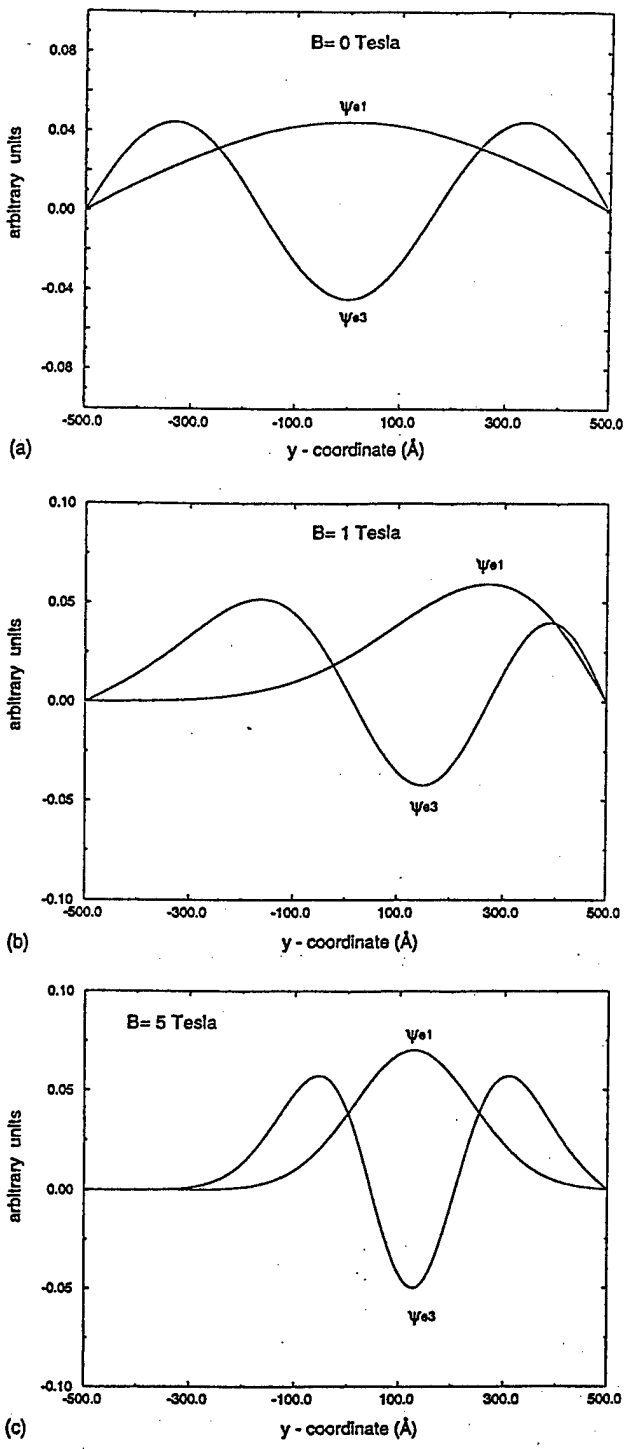



FIG. 6. The $y$ component of the electron envelope functions for the first and third electronic subbands. The results are shown for cases when (a) no magnetic field is present, (b) when a weak magnetic field is present and, finally, (c) when a strong magnetic field is present.

$$\hbar\omega=\sum_i \pm\hbar\omega_i, \qquad \hbar\vec{k}=\sum_i \pm\hbar\vec{k}_i. \quad (10)$$

It is obvious that the third-ranked tensor $\chi^{(2)}$ will vanish in any structure with inversion symmtery. A quantum con-

fined structure may lack inversion symmetry for two main reasons. (i) The semiconductor material by its intrinsic chemical and crystalline structure may lack inversion symmetry,[9] and this is the case in most III–V, II–VI, and I–VII compounds along certain crystallographic directions. (ii) The quantum confining potential well may be asymmetric (e.g. triangular potential well, asymmetric double square well potential, etc.). In the first case, the asymmetry is related to the intracell charge asymmetry and is not affected by the confinement since the latter extends over several unit cells. In the second case, the asymmetry is artificially imposed and therefore can be engineered. It clearly depends on the confining potential and hence an applied electric field can alter the potential and change the degree of symmetry breaking.

In the present work we restrict ourselves to the second case and do not consider intrinsic second-order nonlinearities of GaAs which are actually quite large (the nonlinear susceptibility of bulk GaAs is $\chi^{(2)}_{14}=3.8\times10^{-10}$ m/V[10]). As mentioned before, we avoid an electric field since it promotes carrier escape and we consider a magnetic field instead. Although a magnetic field does not affect the potential to first order, the simultaneous action of *symmetric* electrostatic potential and an external magnetic field may lead to the *uneven charge distribution* along the width ($y$ axis) of the wire caused by different degrees of skewing of the wave functions. Because of this reason, it is possible to break the inversion symmetry in a symmetric quantum well or wire with a magnetic field alone. This approach is superior to applying a transverse electric field since the latter will tilt the confining potential wells thereby promoting carrier escape from the well by either tunneling or thermionic emission.

The large magnitude of the dipole moments associated with otherwise forbidden transitions between subbands of the same parity and their sensitivity to the biasing magnetic field opens up the possibility of second-harmonic generation (SHG) that can be controlled by the magnetic field. In order to evaluate the magnitude and dependences of SHG on the biasing field and wire geometry, we calculate the second-order susceptibility using the formula[11]

$$\chi^{(2)}_{\mu\alpha\beta}(-\omega_\sigma;\omega_1;\omega_2)=\frac{Ne^3}{\epsilon_0 2\hbar^2}\hat{S}_T\sum_{abc}\rho_o(a)\times\left[\frac{d^\mu_{ab}d^\alpha_{bc}d^\beta_{ca}}{(\Omega_{ba}-\omega_1-\omega_2)(\Omega_{ca}-\omega_2)}\right], \quad (11)$$

where $N$ is the concentration (number density) of conduction electrons, $\hbar\Omega_{\alpha\beta}\equiv\hbar\Omega_{\alpha\beta}(B,W)$ is the energy spacing between $\alpha,\beta$ subbands that depends on the applied magnetic field and wire width, $d_{mn}\equiv d_{mn}(B,W)$ is a dipole element calculated using Eq. (4), and $\omega_\sigma$ is defined to be $\omega_\sigma=\omega_1+\omega_2$. The total symmetrization operation $\hat{S}_T$ indicates that the expression that follows it is to be summed over all six permutations of the pairs $(\mu,-\omega_\sigma)$, $(\alpha,\omega_1)$, $(\beta,\omega_2)$. Since $\hat{S}_T$ involves a summation over all possible permutations, it is clear that $\chi^{(2)}_{\mu\alpha\beta}(-\omega_\sigma;\omega_1;\omega_2)$ is invariant under any of them. For simplicity, the Fermi distribution $\rho_o(a)$ was assumed to be unity.

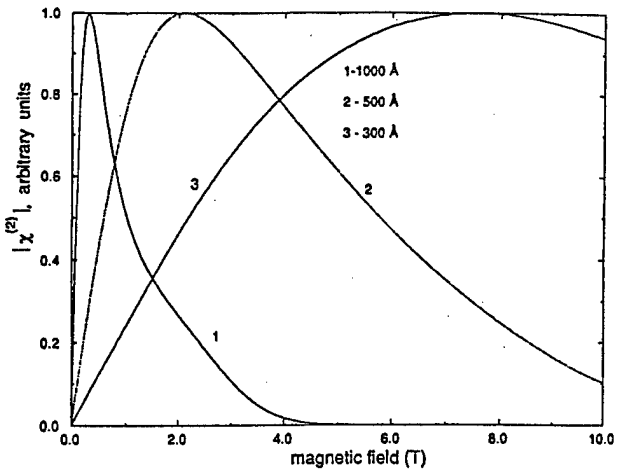



FIG. 7. Second-order susceptibility as a function of the biasing magneti field. The peak values of the susceptibility are $13.2\times10^{-7}$ m/V, 1.: $\times10^{-7}$ m/V and $3\times10^{-8}$ m/V for wire widths of 1000 Å, 500 Å, and 30( Å, respectively. The results are shown for the wave vector $k=0.01$/Å (fixe( excitation frequency).

Eq. (11) is an approximation that applies only under the condition that all of the optical frequencies involved (operational frequencies $\omega_\sigma,\omega_1,\omega_2$) are removed far enough from the subband transition frequencies. It means that the medium is assumed to be transparent and loss free at all the relevan optical frequencies. This assumption can be relaxed by the introduction of transition damping factors into the expressior in Eq. (11). In our study we are mainly interested in the effects of an applied magnetic field on the second-order susceptibility. Since these effects manifest themselves in Eq (11) primarily via the magnetic field dependence of the dipole elements $d_{mn}\equiv d_{mn}(B)$, we did not include any damping constants and associated finite linewidths of the electronic states. One should also note here that Eq. (11) is strictly correct only for dilute media. In this case, one can write $\chi^{(2)}=N\alpha^{(2)}$ with $\alpha^{(2)}$ being the second-order nonlinear polarization. The above expression is valid only under moderate excitation.

In Fig. 7, we present normalized values of $\chi^{(2)}$ as a function of magnetic field for three different wire widths and a fixed value of the wave vector $k$ (fixed excitation frequency). The operational frequencies $\omega_1=\omega_2$ are chosen for a $CO_2$ laser. For wide ranges of magnetic flux densities ($B<20$ T) and wire widths (100 Å $<W<$1000 Å), these frequencies are removed far enough from the subband transition frequencies $\Omega_{\alpha\beta}(B,W)$. As long as the latter is true the $\chi^{(2)}$ dependence on magnetic field is governed mainly by dipole elements $d_{mn}$. Consequently, the $\chi^{(2)}$ curve for $W=1000$ Å peaks at the same value of a magnetic flux density ($B=0.3$ T) as the $e1$-$e3$ dipole curve of Fig. 3. The magnetic flux density at which $\chi^{(2)}$ reaches its maximum increases with decreasing wire width. This happens because it takes a higher magnetic field to condense electronic states into cyclotron orbits (Landau condensation) when the electrostatic confinement is stronger (narrower wires).

Fig. 8 shows the dependence of the normalized values o $\chi^{(2)}$ on wire width for three different values of a magnetic

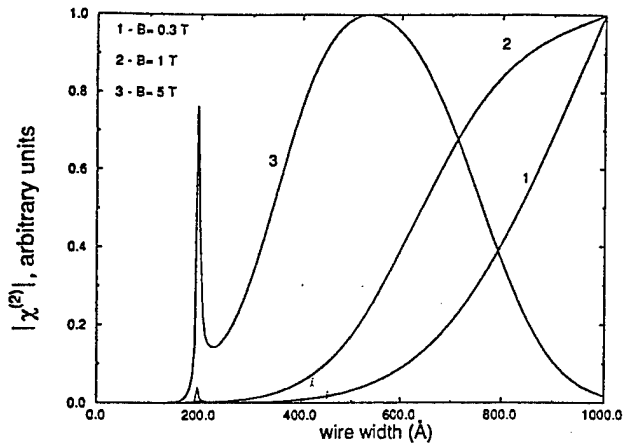



FIG. 8. Second-order susceptibility as a function of the wire width for three values of the biasing magnetic field. The maximum values of the $\chi^{(2)}$ curves are the same as in Fig. 7. The narrow peaks at a wire width of ~200 Å are due to resonances occurring when either $\Omega_{ba}=\omega_1+\omega_2$ or $\Omega_{ca}=\omega_2$.

field and a fixed value of the wave vector. For weak magnetic field of 0.3 T, the $\chi^{(2)}$ curve increases monotonically with increasing wire width. This happens because $d_{m,n}$ is proportional to the wire width $W$ (see Eq. (5) which is valid at zero field). The magnetic field is obviously not strong enough for the onset of Landau condensation. At a moderate magnetic flux density of 1 T, we can observe some saturation features, and for a strong magnetic field of 5 T, the curve is nonmonotonic, rolling down to almost zero for the wire width of 1000 Å. The physics underlying the difference in the behaviors of the three curves is essentially the same as that responsible for the features in Fig. 7. At small values of wire width ($W\approx 170$ Å), there is an additional peak in the $\chi^{(2)}$ curve. This peak is a manifestation of the fact that $\Omega_{ba}$ has become comparable to the operational frequencies, i.e. $\Omega_{ba}\approx\omega_1+\omega_2$.

In our numerical calculations we have used $N=10^{17}$ $cm^{-3}$. For this dilute concentration, high density effects such as screening and bandgap renormalization are not important and Eq. (11) is strictly valid. In fact, Ref. 1 demonstrated excellent agreement between theory and experiment without accounting for any high density effect even though the carrier concentration in that study was $N=5\times10^{17}/cm^3$. Therefore, we believe that high density effects are not significant in this regime.

The peak value of the second-order susceptibility for a wire width of 500 Å is $\chi^{(2)}=1.5\times10^{-7}$ m/V (the absolute magnitudes of the peak values for various wire widths are given in the caption of Fig. 7). For comparison, the nonlinear susceptibility of electric field biased GaAs quantum wells ($W=92$ Å), calculated theoretically and measured experimentally in Ref. 1, was $\chi^{(2)}=2.4\times10^{-8}$ m/V for an electric field of 36 kV/cm. This shows that relatively weak magnetic fields in quantum wires can produce similar magnitudes of $\chi^{(2)}$ as rather strong electric fields in quantum wells. Unfortunately, to our knowledge, there is no theoretical or experimental result available for either electric field biased quantum wires or magnetic field biased quantum wells so a direct comparison is not possible. Nonetheless, it is obvious that magnetic field biased quantum wires provide a very attractive alternative to other means of producing large $\chi^{(2)}$ values. In fact, the largest value of $\chi^{(2)}$ (obtained at a magnetic flux density of 2 T) in a magnetic-field-biased quantum wire is found to be three orders of magnitude higher than what can be achieved in bulk GaAs.

## IV. CONCLUSION

We have theoretically studied the giant dipole effect in magnetic-field-biased semiconductor quantum wires. The dipoles are associated with transitions between magneto-electric subbands within the conduction band, some of which are forbidden in the absence of the magnetic field. The resonant frequencies of these transitions can be tuned by the magnetic field which allows the realization of externally tunable intersubband lasers. We have also studied the possibility of second-harmonic generation in a quantum wire biased with a magnetic field and find a strong second-harmonic component of the susceptibility. This may have important applications in nonlinear optics.

## ACKNOWLEDGMENT

This work is supported by the U.S. Army Research Office under Contract Nos. DAAH04-95-1-0586 and DAAH04-95-1-0527.

[1] M. M. Fejer, S. J. B. Yoo, R. L. Byer, A. Harwit, and J. S. Harris, Phys. Rev. Lett. **62**, 1041 (1989), and references therein; L. C. West and S. J. Eglash, Appl. Phys. Lett. **46**, 1156 (1985); for an overview of current work on intersubband transitions, see, for example, *Quantum Well Intersubband Transition Physics and Devices*, NATO ASI Series E270, edited by H. C. Liu, B. F. Levine, and S. Y. Anderson (Academic, Dordrecht, 1994).

[2] A. Sa'ar, I. Grave, N. Kuze, and A. Yariv, *Nonlinear Optics: Materials, Phenomena and Devices*, (IEEE, New York, 1990) 113; B. F. Levine, R. J. Malik, J. Walker, K. K. Choi, C. G. Bethea, D. A. Kleinman, and J. M. Vandenberg, Appl. Phys. Lett. **50**, 273 (1987).

[3] See, for example, Manfred Helm, Semicond. Sci. Technol. **10**, 557 (1995).

[4] E. J. Austin and M. Jaros, Phys. Rev. B **31**, 5569 (1985).

[5] A. Balandin and S. Bandyopadhyay, J. Appl. Phys. **77**, 5924 (1995); A. Balandin, Ph.D. dissertation, University of Notre Dame, 1996.

[6] C. Weisbuch and B. Vinter, *Quantum Semiconductor Structures: Fundamentals and Applications* (Academic, Boston, 1991).

[7] S. Chaudhuri and S. Bandyopadhyay, J. Appl. Phys. **71**, 3027 (1992).

[8] See, for example, J. M. Hvam in *Nonlinear Spectroscopy of Solids: Advances and Applications*, edited by B. Di Bartolo and B. Bowlby (Plenum, New York, 1994), pp. 91–149.

[9] This lack of symmetry is due to the transfer of a valency charge from one atom to the other bond-forming atom and resultant uneven charge distribution along the bond axis.

[10] *Handbook of Lasers*, edited by R. J. Pressley (Chemical Rubber, Cleveland, 1971), p. 504.

[11] See, for example, P.N. Butcher and D. Cotter, *The Elements of Nonlinear Optics* (Cambridge University Press, Cambridge, 1990); A. Yariv, *Quantum Electronics* (Wiley, New York, 1989); or Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley, New York, 1984).

Presented at the Tenth International Conference on Superlattices, Microstructures and Microdevices, July 8-11, 1997, Lincoln, Nebraska

# Intrinsic Optical Non-linearity and Second Harmonic Generation in Electrochemically Self-Assembled CdS Quantum Dots

A. Balandin, S. Bandyopadhyay and P. G. Snyder
Department of Electrical Engineering, University of Nebraska
Lincoln, Nebraska 68588 USA

S. Stefanovich, A. Varfolomeev, D. Zaretsky
Russian Research Center "Kurchatov Institute"
Kurchatov Square, Moscow 123182, Russia

G. Banerjee and A. E. Miller
Department of Chemical Engineering, University of Notre Dame
Notre Dame, Indiana 46556, USA

**Abstract**

We report the observation of strong second harmonic optical signal generation in 15-nm sized CdS quantum dot arrays produced by electrochemical self-assembly. At incident photon energies well below the bandgap of CdS, the measured second-order susceptibility $\chi^{(2)}$ is five times larger than that of bulk CdS. Some speculations about the origin of this enhanced second-order optical non-linearity are presented.

## 1 Introduction

Optical non-linearities causing frequency conversion are useful in a number of applications such as mixing, switching, limiting and coupling. Most ordinary solids, however, are not efficient frequency converters because they exhibit extremely small higher-order components of dielectric susceptibilties, and also because phase matching, which optimizes frequency conversion, is difficult to attain in solids that are not birefringent. Furthermore, if the solid has inversion symmetry, then it can exhibit no even-order susceptibility[1] unless the symmetry is broken artificially either by the intentional growth of an asymmetric structure[2], or by an external field; electric[3] or magnetic.[4]. On rare occasions, inversion symmetry can be broken spontaneously by built-in fields.

Typically, second-order optical non-linearity in a solid has two sources; a bulk contribution and a surface or interface contribution.[5] The inversion symmetry is automatically broken at an interface because of the discontinuity of the crystalline structure and the large gradient of the normal-to-surface component of the electric field of incident radiation. This alone can result in finite (but small) second-order dielectric susceptibility $\chi^{(2)}$ which causes weak second harmonic generation (SHG). Other possible mechanisms of SHG in semiconductors include (i) electric quadrupole nonlinearity,[6] (ii) deformations and stresses in the structure leading to

internal fields and nonlinear polarization, (iii) lowering of the symmetry of the crystal under the action of the intense electric field of incident radiation, and (iv) high degree of disorder of the crystalline structure (quasi-amorphism).

In this paper, we report the preliminary observation of second harmonic generation in ordered regimented arrays of 15-nm diameter CdS quantum dots that were electrochemically self-assembled. We believe that all of the above mechanisms play some role in SHG in these samples. However, regardless of its origin, the very observation of SHG in these samples is promising because it provides a cheap and convenient way to produce non-linear optical components. These structures may find use in frequency converters, mixing and limiting.

The rest of the paper is organized as follows. In section 2, we describe sample preparation. Section 3 outlines experimental setup and the measurement technique. In section 4, we present results and estimate the second-order susceptibility. Conclusions are given in section 5 of the paper.

## 2 Sample Preparation and Characterization

Our samples are electrochemically self-assembled quasi-periodic two-dimensional arrays of CdS quantum dots embedded in alumina. The average *physical* diameter of the dots 15 nm, but the effective *optical* diameter is significantly less because of side depletion. They are synthesized by first anodizing a thin foil of aluminum for a few seconds in 15% $H_2SO_4$ at a dc current density of 40 mA/cm$^2$ to produce a nanoporous alumina film on the surface of the foil. These pores form an ordered hexagonal array with an average diameter of 15 nm with a 7% standard deviation. A bright field TEM image of a typical pore assembly is shown in Fig. 1. The pores are then selectively filled up with CdS using ac electrodeposition. This is accomplished by immersing the alumina film in a boiling aqueous solution of $CdSO_4$. An ac potential of 18 volts rms is applied for different durations (5-30 seconds) using the sample and a graphite rod as electrodes. The $Cd^{++}$ ion in the solution reacts with the $S^{--}$ ion left behind in the walls of the pores from the previous anodization step to form CdS. The amount of CdS produced (pore filling) increases superlinearly with the duration of electrodeposition. The pores have been directly imaged with TEM and field-emission SEM[7]. Presence of material in the pores has been verified with cross-section TEM and stoichiometry was checked with Auger. Optical signatures of CdS were also verified in the past by Raman, photoluminescence, reflection and absorption[7,8]. These measurements revealed strong quantum confinement effects. In fact, the measured blue-shift in the optical sepctra indicates that the effective optical diameter of the dots is $\sim$ 3.5 nm (even though the physical diameter is 15 nm)[7,8]. The size reduction is presumably caused by side depletion around the periphery of the dots due to interface states. These are some of the smallest quantum dots that have shown optical activity.

We have also performed variable angle spectroscopic ellipsometry study of the samples[9] in order to determine (i) the thickness of the alumina layer in which the quantum dots are embedded, (ii) the shape of the CdS dots, (iii) the volume fraction of CdS in the alumina

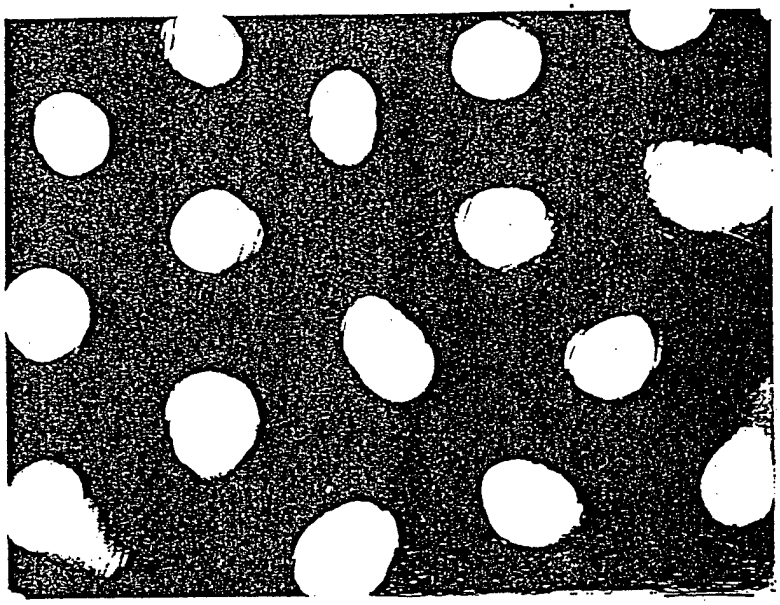


Figure 1: Bright field TEM image of a porous alumina film showing a regimented array of nanopores. The average diameter of the pores is 15 nm with a 7% standard deviation. These pores are filled up with CdS by electrodeposition to create the CdS quantum dot arrays.

layer, and (iv) surface roughness and lateral thickness non-uniformity. To obtain information about sample optical constants and structure, ellipsometry data were fitted with a multilayer model based on effective medium approximation. Model parameters were varied (using the Levenberg-Marquardt algorithm) to minimize the mean-square error between two sets of data: measured and model-generated. The samples were viewed as arrays of CdS dots in an alumina matrix that could be treated within the Maxwell-Garnett formalism. This allowed us to determine the volume fraction of CdS in the alumina layer, and the depolarization factor directly related to dot shape. It was found that the CdS dots can be best modeled by rotational ellipsoids with a depolarization factor $\approx 0.7$. The thickness non-uniformity is about 7%; and the volume fraction of CdS for 10 second deposition time was found to be also $\sim 7\%$.

## 3 Experimental Setup and Measurements

For optical SHG measurements, we used a mode-locked Nd:YAG laser producing 10-ns-long pulses with an average power density of 1 W/cm$^2$ and absolute peak intensity of about 1 $MW/cm^2$ at a wavelength of 1.064 $\mu m$ (photon energy = 1.32 eV). This corresponds to sub-bandgap radiation. The bulk bandgap of CdS is 2.4 eV and the optically measured bandgap, enhanced by quantum confinement, is about 3 eV for samples prepared by 10 seconds electrodeposition[7]. The samples were irradiated by laser light at an angle of 20

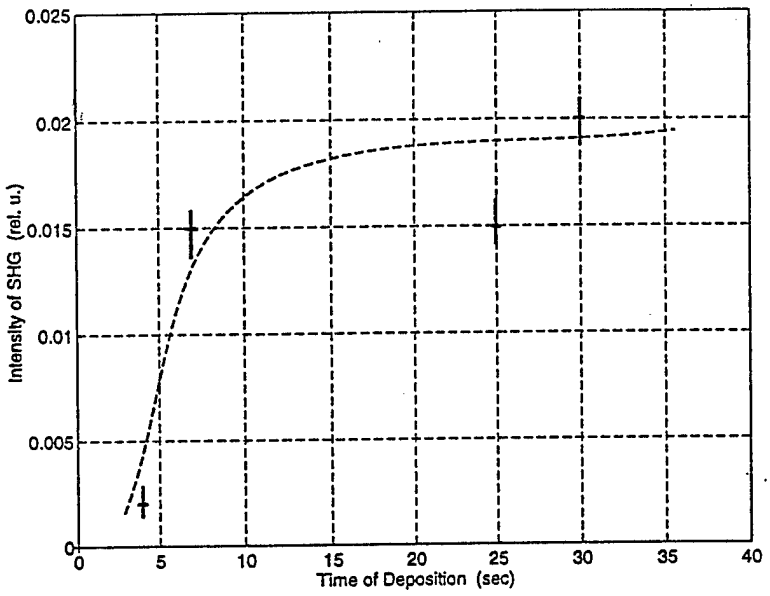



Figure 2: The intensity of the SHG in the arrays of CdS quantum dots relative to the intensity of SHG in the quartz etalon. Time of deposition determines the quantity of CdS in the alumina pores.

degrees and the spotsize was about 1cm which is approximately equal to the sample size. The second harmonic signal was observed in the reflection mode by collecting all radiation reflected back (in $\pi$ rad).

A quartz ($\alpha - SiO_2$) etalon was used as a reference. The nonlinear coefficient of the etalon is $d_{11} = 4.4 \times 10^{-13}\ m/V$ and the refractive indices, in the spectral region of interest, are about $n_o = 1.5350$ and $n_e = 1.5438$ for ordinary and extraordinary rays, respectively. The measured absolute intensity of the SHG of the quartz etalon was about $10^{-15}\ W/cm^2$.

In Fig. 2, we show the normalized intensity of the SHG in arrays of CdS quantum dots as a function of the time of deposition. As mentioned before, the time of deposition determines the height of CdS ellipsoids in the alumina pores. The maximum value of the second harmonic intensity, $I(2\omega)$, for our sample with the highest volume fraction of $CdS$ is about 2% of the $I(2\omega)$ of the quartz etalon.

## 4 Results

In order to characterize the process of SHG quantitatively, we use the following formula for the conversion efficiency[10]

$$\frac{I(2\omega)}{I(\omega)} = \frac{2\omega^2 |d_{eff}|^2 l^2 I(\omega)}{n^3 c^3 \epsilon_o} sinc^2(\vec{\Delta k}\vec{l}/2), \qquad (1)$$

where $sinc(x) \equiv sin(x)/x$, $I(\omega)$ is the intensity of incident light at the fundamental frequency, $l$ is the interaction length, $d_{eff}$ is the effective material nonlinear coefficient, $n$ is the refractive index of the material, $c$ is the speed of light, and $\Delta\vec{k} = \vec{k}(2\omega) - 2\vec{k}(\omega)$ is the wave vector mismatch. Phase matching determines the phase synchronism factor $sinc^2(\Delta\vec{k}\vec{l}/2)$ which is unity at $\Delta\vec{k}\vec{l} = 0$. The above formula is valid in the limit of small conversion and plane wave focusing. It is reasonable to assume that the interaction length, $l$, in our case is approximately equal to the size of a quantum dot. Since the size of the dot is very small (vertical dimension of the dots is of the order of 50$\mathring{A}$) and the difference between refractive indices at the fundamental and doubled frequencies is insignificant, we can limit our consideration to the case when $\Delta\vec{k}\vec{l} \approx 0$. Under such conditions, one can estimate from Eq. (1) the value of the effective nonlinear coefficient $d_{eff}$. We find that $d_{eff} = 8.8\times10^{-11}$ m/V which is five times larger than that of bulk CdS. Since $d_{eff}$ is proportional to the second-order susceptibility $\chi^{(2)}$, we conclude that our quantum dots have a value of $\chi^{(2)}$ that is five times larger than that of bulk medium. We should point out that the overall $\chi^{(2)}$ of our samples is actually much smaller since only 6-7% of the samples consist of CdS dots. Taking this factor into account, the average $d_{eff}$ for our samples is $5.28\times10^{-12}$ m/V.

## 5 Discussion and Conclusion

It should be noted that the second-order non-linearity in our samples is observed at below-bandgap frequency. This indicates that the non-linearity is either related to virtual proccesses or intraband processes. As far as the latter is concerned, the second harmonic generation is may be associated with inter-subband dipole transitions in the conduction band which is a well-known mechanism[4]. In our samples, the calculated energy spacing between the electronic subbands is between 1 and 1.5 eV (assuming infinite potential barriers at the interfaces and typical amount of side depletion) while the incident photon energy is 1.32 eV. Consequently, the incident radiation may be exciting real transitions between subbands. This mechanism is obviously directly related to quantum confinement and would not be relevant for bulk samples. Other possible origins of the SHG include electric quadrupole nonlinearity,[6] deformations and stresses in the structure leading to internal fields and nonlinear polarization, lowering of the symmetry of the crystal under the action of the intense electric field of the incident radiation, and high degree of disorder of the crystalline structure (amorphism).

An interesting question is what breaks the inversion symmetry in these structures to cause even order non-linearities. The obvious answer is local electric fields associated with trapped charge centers in the alumina matrix. Local fields can also be generated by stresses and dislocations.

In conclusion, we have shown that electrochemically self-assembled semiconductor quantum dot arrays can exhibit strong second-order non-linearities in their optical properties. The magnitude of the second-order susceptibility is five times larger than that of bulk CdS. Polarization studies and study of the dependence of SHG on the angle of incidence are planned for the near future to further elucidate the origin of this enhancement.

## Acknowledgements

This work was supported by the US Army Research Office under grant DAAH04-95-1-0586.

## References

[1] See, for example, P.N. Butcher, D. Cotter, *The Elements of Nonlinear Optics,* (Cambridge University Press, Cambridge, 1990); A. Yariv, *Quantum Electronics,* (Wiley, New York, 1989); Y.R. Shen, *The Principles of Nonlinear Optics,* (Wiley, New York, 1984).

[2]. A. Sa'ar, I. Grave, N. Kuze, and Yariv in *Nonlinear Optics: Materials, Phenomena and Devices* (IEEE, New York, 1990) 113.

[3]. M.M. Fejer, S.J.B. Yoo, R.L.Byer, A. Harwit, J.S. Harris, *Phys. Rev. Lett.*, **62**, 1041 (1989).

[4]. A. Svizhenko, A. Balandin and S. Bandyopadhyay, *J. Appl. Phys.*, **81**, 7927 (1997); A. Balandin, A. Svizhenko, S. Bandyopadhyay, in *Nanostructures: Physics and Applications*, Eds. Zh. Alferov and R. Suris, (Russian Academy of Sciencies, St. Petersburg, Russia, 1997), 294.

[5]. O.A. Aktsipetrov, A.A. Fedyanin, A.V. Melnikov, J.I. Dadap, X.F. Hu, M.H. Anderson, M.C. Downer, J.K. Lowell, *Thin Solid Films*, 294, 231 (1997).

[6]. V.N. Denisov, B.N. Mavrin, V.B. Podobedov, Kh. Sterin, B.G. Varshal, *Opt. Spectrosc. (USSR)*, 49, 221 (1980).

[7]. S. Bandyopadhyay, et. al., *Nanotechnology*, Vol. 7, 360 (1996); R.E. Ricker, A.E. Miller, D.-F. Yue, G. Banerjee, and S. Bandyopadhyay, *J. Electron. Materials*, Vol. 25, No. 10, 1585 (1996).

[8]. S. Bandyopadhyay, A. E. Miller and M. Chandrasekhar, in *Proc SPIE*, Vol. 2397, (1995), pp. 11-30; D-F. Yue, G. Banerjee, A.E. Miller, S. Bandyopadhyay, *Superlatt. and Microstr.*, Vol. 18, No. 1, 702 (1995).

[9]. A. Balandin, PhD Dissertation, University of Notre Dame, USA, 1996.

[10]. R. L. Bayer, in *Non-linear Optics*, Eds. P. G. Harper and B. S. Wherrett (Academic Press, New York, 1977), p. 61.

Presented at the Fourth International Symposium on Quantum Confinement: Nanoscale Materials and Devices, 191st Meeting of the Electrochemical Society, Montreal, Canada, July 4-9, 1997

# Infrared Second Harmonic Generation in Magnetic-Field-Biased Quantum Wires

A. Svizhenko, A. Balandin, and S. Bandyopadhyay
Department of Electrical Engineering
University of Nebraska
Lincoln, Nebraska 68588

**Abstract**

We have theoretically studied non-linear frequency conversion in a semiconductor quantum wire biased with a magnetic field. In these systems, efficient second harmonic generation occurs as a result of the large value of the second-order dielectric susceptibility $\chi^{(2)}$ arising from dipole transitions between magneto-electric subbands. The magnitude and peak frequency of $\chi^{(2)}$, as well as the absorption and refractive index associated with $\chi^{(2)}$, can be tuned with the magnetic field. This allows one to achieve low insertion loss and efficient phase matching by manipulating the absorption and refractive index with a magnetic field.

## I. Introduction

Most ordinary solids are not efficient frequency converters because they exhibit extremely small even-order dielectric susceptibilties. Ideally, even-order susceptibilities vanish in solids with inversion symmetry.[1] Consequently, a semiconductor structure can exhibit a large value of the second order susceptibility $\chi^{(2)}$ only if the inversion symmetry of the conduction-band potential is broken artificially either by an external electric field, or by the intentional growth of an asymmetric structure. Obviously, the former is the preferred method since an electric field can be varied continuously and

this allows one to tune the degree of symmetry-breaking and the magnitude of $\chi^{(2)}$. However, an electric field has a practical shortcoming. In a quantum confined structure, it tilts the potential barriers confining the photogenerated carriers. As a result, carriers can escape by tunneling or thermionic emission and this is especially serious in *GaAs/AlGaAs* systems where the barrier height is relatively small. Indeed, it has been pointed out that the electronic states in a quantum confined system biased by a transverse electric field are never true bound states since the particles can always lower their energy by escaping from the well[2]. Consequently, optical transitions (and their higher order harmonics) associated with these states have inconveniently large linewidths and small oscillator strengths.

Recently, we proposed magnetostatic biasing as an attractive alternative to mitigate this problem.[3] We showed that a magnetic field can break inversion symmetry in a quantum wire without tilting potential barriers. A transverse magnetic field, applied to a wire, exerts a Lorentz force on an electron moving along the length. As a result, its wave function (in any magneto-electric subband) will be skewed towards one edge of the wire. This skewing does not tilt potential barriers to first order (the barriers may tilt slightly because of a second-order effect associated with space-charges and the self-consistent (Hall) electric field). However, it effectively breaks inversion symmetry since it causes net charges to accumulate at either edge of the wire. This leads to a non-vanishing even-order susceptibility in an otherwise symmetric structure. The skewing has another subtle effect. The degree to which the wave function is skewed is *different in different subbands* since an electron has different kinetic energies (and hence experiences different Lorentz forces) in different subbands. As a result, transitions between subbands whose wave functions have the same parity - which are forbidden without a magnetic field - are now allowed since the parities are altered by different amounts in different subbands. This effect has some similarity with the quantum confined Lorentz effect (QCLE) previously examined by us[4] in

the context of interband transitions between conduction and valence band states.

In this paper, we first calculate the second-order susceptibility $\chi^{(2)}$ in a symmetric quantum wire whose inversion symmetry (along the width) has been broken with a magnetic field. We restrict ourselves to narrow GaAs wires with a width of about 150$\mathring{A}$. The energy spacing between the first and second subband is $\Delta E_{12} \approx 72$ meV. This choice of the wire dimension puts the resonant frequency (for transitions between the lowest subbands) in the mid-infrared spectral region. The wavelength of the second harmonic component of this transition is about 8.6 $\mu$m. Here we will be mainly interested in $\chi^{(2)}$ arising from resonant and near-resonant inter-subband transitions which are governed by the interplay of dipoles and resonant excitations. In contrast, Ref. [3] focussed on the off-resonance regime which was governed solely by the dipoles. We will also calculate absorption and refractive index in the frequency region of interest for both pump and second harmonic frequencies and show how they can be manipulated with an external magnetic field to realize low insertion loss and efficient phase matching.

The rest of the paper is organized as follows. In the next section, we describe the theoretical formulation, followed by results. Finally, in section IV, we present the conclusions.

## II. Theory

We consider a generic *GaAs* quantum wire (as shown in the inset of Fig. 1) with a magnetic field applied along the $z$ direction. The thickness along the $z$ direction is so small (and consequently the subband separation in energy in this direction is so large) that for the range of photon energies considered, an electron cannot be excited (by real transition) into a subband which has more than two nodes along the z-direction. In other words, such a transition will not be accessible in energy. This restriction, coupled with the fact that a magnetic field does not affect the z-component of the electron wave function, allows us to drop the z-component from further consideration. The width

of the wire along the y-direction is however large enough ($W = 150\mathring{A}$) that subbands with more than two nodes along the y-direction are accessible in energy.

In systems without inversion symmetry, the lowest order optical nonlinearity is of the second order and is expressed by

$$\vec{P^{(2)}}(\vec{k},\omega) = \chi^{(2)}(\omega;\omega_1,\omega_2)\vec{E_1}(\vec{k_1},\omega_1)\vec{E_2}(\vec{k_2},\omega_2), \qquad (1)$$

where $\vec{P}$ is the polarization caused by two electric fields $\vec{E_1}$ and $\vec{E_2}$ that are associated with the electromagnetic fields of either two frequency components of the same light beam or two different coherent beams with frequencies $\omega_i$ and wave vectors $\vec{k_i}$. It is well known that the third-ranked tensor $\chi^{(2)}$ will vanish in any structure with inversion symmetry. A quantum confined structure may lack inversion symmetry for two main reasons. (i) the semiconductor material by its intrinsic chemical and crystalline structure may lack inversion symmetry,[5] and this is the case in most III-V, II-VI, and I-VII compounds along certain crystallographic directions, or (ii) the quantum confining potential well may be asymmetric (e. g. triangular potential well, asymmetric double square well potential, etc.). In the first case, the asymmetry is related to the intracell charge asymmetry and is not affected by the confinement since the latter extends over several unit cells. In the second case, the asymmetry is artificially imposed and therefore can be engineered. It clearly depends on the confining potential. Insofar as an applied electric field can alter the potential, it can change the degree of symmetry-breaking and hence modulate $\chi^{(2)}$.

In the present work we restrict ourselves to the second case and do not consider intrinsic second order nonlinearities which can be quite large in some materials (the nonlinear susceptibility of bulk GaAs is $\chi^{(2)}_{14}$=3.8 $10^{-10}$ m/V).[6] As mentioned before, we avoid the use of a symmetry-breaking electric field since it promotes carrier escape. Instead, we consider a magnetic field. Although a magnetic field does not directly affect the potential, it leads to an *uneven charge distribution* along the width ($y - axis$) of the wire because of

the different degrees of skewing of the wave functions in different magneto-electric subbands. This has the effect of breaking inversion symmetry.

As mentioned before, a magnetic field induces forbidden transitions between subbands of the same parity. The large magnitude of the dipoles associated with these transitions and their extreme sensitivity to the field open up the possibility of *controllable* second harmonic generation (SHG) that can be manipulated by the magnetic field. In order to evaluate the magnitude and dependence of SHG on the biasing field and wire geometry, we calculate the second order susceptibility using the formula[1]

$$\chi^{(2)}(2\omega;\omega;\omega) = \frac{Ne^3}{\epsilon\epsilon_o 2\hbar^2}\hat{S}_T \sum_{abc} \frac{d^{\mu}_{ab} d^{\alpha}_{bc} d^{\beta}_{ca}}{(\Omega_{ba} - \omega - i\gamma)(\Omega_{ca} - 2\omega - i\gamma)}, \qquad (2)$$

where $N$ is concentration (number density) of conduction electrons, $\gamma$ is a damping factor associated with elastic and inelastic scattering, $\hbar\Omega_{ba} \equiv \hbar\Omega_{ba}(B, W, k)$ is the energy spacing between the b-th and a-th magnetoelectric subbands which depends on the applied magnetic field, wire width and electron wave vector and $d_{mn} \equiv d_{mn}(B, W, k)$ is a dipole element of transitions between different subbands. The total symmetrisation operation $\hat{S}_T$ indicates that the expression which follows is to be summed over all permutations of the pairs $(\mu, 2\omega), (\alpha, \omega), (\beta, \omega)$. Since $\hat{S}_T$ involves a summation over all possible permutations, it is clear that $\chi^{(2)}_{\mu\alpha\beta}(2\omega;\omega;\omega)$ is invariant under any of them.

In order to calculate dipole elements $d_{mn}(B, W)$, we proceed as in Ref. [3]. Under the electric dipole approximation, the matrix element of photoinduced inter-subband transitions within the conduction band is given by[7]

$$d_{f,i}(k, B) = e \int \chi_f(y, k, B)\hat{\eta} \cdot \vec{r}\chi_i(y, k, B)d\vec{r} \int u^*_f(x, y, k)u_i(x, y, k)d\Omega \qquad (3)$$

where $d\Omega$ is a volume element, $\hat{\eta}$ is the unit vector along the direction of the incident photon polarization, $\vec{r} = x\vec{a}_x + y\vec{a}_y$ is the two-dimensional radius vector, and subscripts $i, f$ stand for initial and final states respectively. Now, if we assume that the incident light is polarized along the y-direction so that

$\hat{\eta} = \hat{a}_y$, the above equation simplifies to

$$d_{f,i}(k,B) = e < \chi_f|y|\chi_i >= e\int_{-W/2}^{W/2} y\chi_f(y,k,B)\chi_i(y,k,B)dy, \qquad (4)$$

where $W$ is the width of the quantum wire along the y-direction. One should note here, that if there is no magnetic (or electric) field applied, the envelope functions $\chi_i$ are just particle-in-box states and the dipole moment in Eq. (4) is non zero only for the transitions between subband states of opposite parity. However, this is obviously not the case when a magnetic field is present. It is clear from Eq. (4) that to calculate the dipole moments in the presence of a magnetic field, all we need to compute are the wave functions $\chi_{f,i}(y,k,B)$ at a given magnetic field $B$, for given magnetoelectric subbands $f$ and $i$, and for a given wave vector $k$. This is achieved via a numerical (finite difference) solution of the Schrödinger equation for the y-component of the wavefunction

$$\frac{\partial^2\chi(y)}{\partial y^2} + \frac{2m^*}{\hbar^2}E\chi(y) - \left(\frac{y}{l^2}\right)^2\chi(y) + 2\frac{y}{l^2}k\chi(y) - k^2\chi(y) = 0 \qquad (5)$$

with $l$ being the magnetic length given by $l = \sqrt{\hbar/eB}$, assuming hardwall boundary conditions

$$\chi(y = W/2) = \chi(y = -W/2) = 0 \qquad (6)$$

and following the prescription of Ref. [8]. Once this is done, we can calculate the dipole moment in Eq. (4) for any chosen intersubband transition, at any chosen magnetic field and for any chosen wave vector.

The absorption of both the fundamental frequency (pump) and its second-order harmonic is very important when considering frequency conversion with *low insertion loss*. In general, it is desirable to have large absorption coefficient $\alpha(\omega)$ for the pump frequency and small $\alpha(2\omega)$ for the converted frequency so that the latter is not re-absorbed to cause large insertion loss. In order to obtain the absorption coefficients in the whole range of frequencies and for different values of a magnetic flux density, we need to calculate the first-order susceptibility as follows

$$\chi^{(1)}(\omega) = \frac{Ne^2}{\epsilon\epsilon_o\hbar}\sum_{ab}\frac{(d_{ab})^2}{\Omega_{ba} - \omega - i\gamma}, \qquad (7)$$

where we have used the same notation as in Eq. (2). The imaginary part of $\chi^{(1)}(\omega)$ is related to the absorption coefficient while the real part is related to the refractive index.

## II. Results

We now present results of our calculations. The physical parameters used for the numerical calculations are relevant for a $GaAs$ quantum wire with relative dielectric constant $\epsilon_r = 12.9$, and effective masses $m_e = 0.067m_o$ and $m_h = 0.5m_o$ where $m_o$ is the free electron mass.

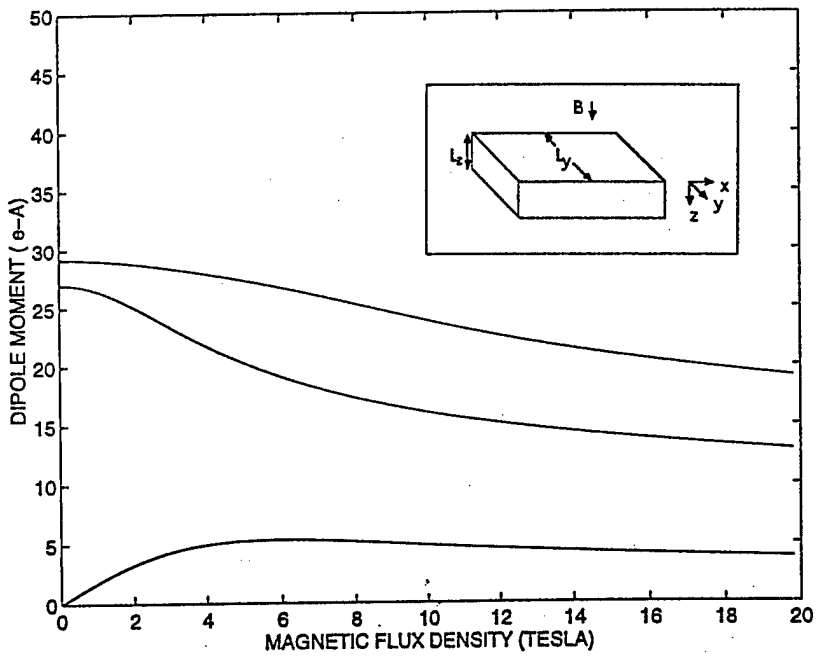



Figure 1: The dipoles of three inter-subband transitions as functions of the applied magnetic field. The induced dipole $d_{e1-e3}$ peaks at a magnetic flux density of 5.3 tesla.

Fig. 1 presents the dipole moments for the lowest intraband transitions as a function of magnetic flux density. At zero magnetic field, a non-vanishing dipole matrix element occurs only for transitions between states of opposite parity (e1-e2, e2-e3) as expected from Eq. (4). Transition dipole $d_{e3-e1}$

has a non-monotonic dependence on the magnetic field. This transition is forbidden at zero field since the wave functions of the first and third subband have the same parity. At low and moderate magnetic fields, the parities are altered by the skewing of the wavefunctions. The skewing effect of the wave functions of the first and the third subbands is shown in Fig. 2 (top). Its degree depends on a subband number, which causes a breaking of inverion symmetry and, consequently, non-zero value of dipole matrix element $d_{e3-e1}$ for otherwise forbidden transition.

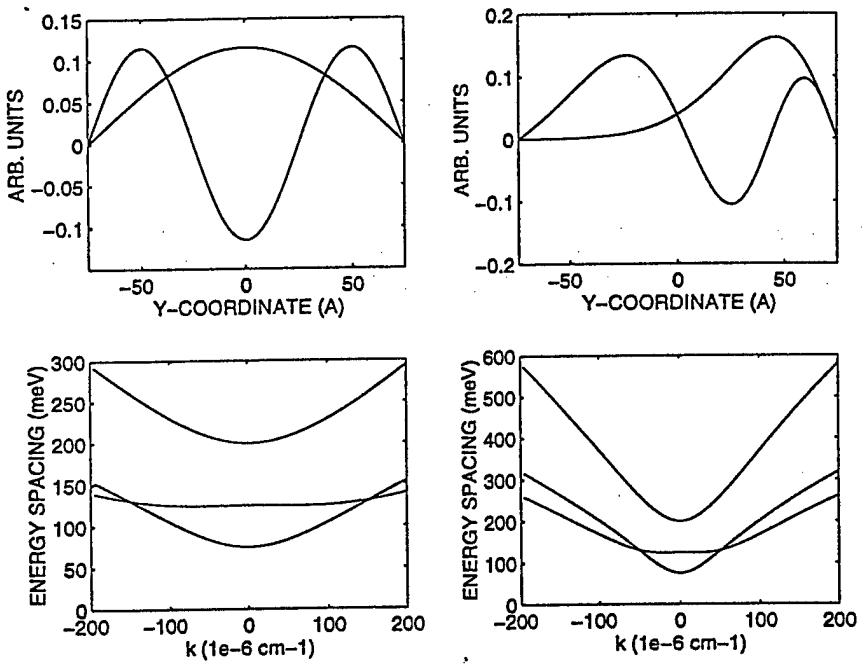



Figure 2: (Top left and right). Skewing of the wave functions of the first and third subbands in a magnetic field. The left panel corresponds to zero magnetic flux density and the wave functions are particle-in-a-box states. The right panel corresponds to a flux density of 3 tesla and the wave functions are those of "edge states". (Bottom left and right). The energy spacing $\hbar\Omega_{mn}$ between the $m$th and $n$th subbands vs. wave vector $k$ at a magnetic flux density B=1 tesla (left) and B=3 tesla (right). The lowest curve (at k=0) corresponds to e1-e2, the intermediate curve to e2-e3, and the highest to e1-e3.

The dipole moment reaches a maximum of about 6 e-Å and then de-

creases. This later decrease is related to the following effect. For some fixed wave vector $k$ (electron velocity), a sufficient increase in the flux density $B$ forces the traversing states ("skipping orbits" or "edge states") to condense into closed cyclotron orbits (Landau levels) which are no longer skewed by the magnetic field to the wire edge since they have no translational velocity and hence experience no Lorentz force. While edge states have a skewed wave function which is not symmetric in space, cyclotron orbits have a wave function that is symmetric about the orbit center. Note that the orbit center's coordinates depend only on $k$ and $B$. Therefore, at a fixed $k$, the wave functions of the first and third Landau levels are symmetric about a *common* center. Whenever this kind of symmetry holds, $d_{e3-e1}$ vanishes. Therefore, the dipole moment $d_{e3-e1}$ decreases gradually to zero at high magnetic field with the onset of Landau condensation (5 tesla for this wire dimensions).

In Fig. 3, we plot the absolute values of $\chi^{(2)}$ as a function of photon energy for two different values of the magnetic field. In our numerical calculations we have used a dilute carrier concentration of $N$=$10^{17}$ $cm^{-3}$ which allows us to neglect high density effects such as screening and bandgap renormalization. Both susceptibility curves have pronounced three-peak resonant structure which corresponds to two one-photon transitions e1-e2 (at 72 meV) and e2-e3 (at 124 meV) and one two-photon transition e1-e3 (at 100 meV) between magneto-electric subbands. These three peaks have different broadenings because the sum in Eq. (2) represents an integral effect of all direct transitions with different values of electron wave vector $k$ and because of the complex dependence of the subband spacing $\hbar\Omega_{mn}$ on $k$ and $B$ (see Fig. 2 (bottom)). The latter also gives rise to an uneveness in the second order susceptibility peaks. The peak value of the second order susceptibility is $\chi^{(2)} = 14.5\mathring{A}/V$ for 1 tesla field (left panel); and $\chi^{(2)} = 43.1\mathring{A}/V$ for 3 tesla (right panel). For comparison, the nonlinear susceptibility of electric field biased GaAs quantum wells (W=92 $\mathring{A}$) - calculated theoretically and measured experimentally in Ref. [9] - was $\chi^{(2)}$=240Å/V for an electric field of 36

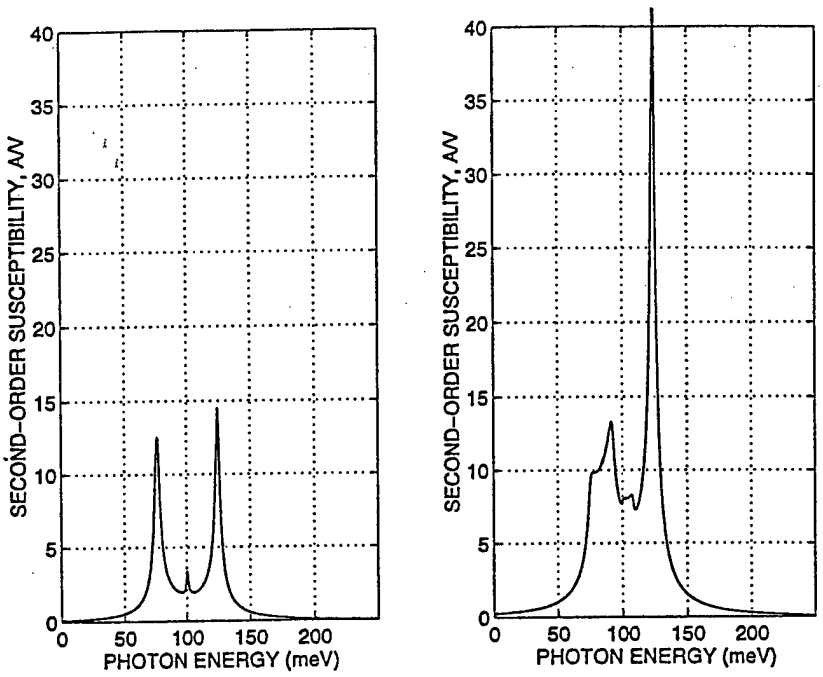



Figure 3: Second order susceptibility as a function of the photon energy for two values of the biasing magnetic field. The left panel corresponds to a magnetic flux density B=1 tesla, the right panel to B=3 tesla. The maximum values of the $\chi^{(2)}$ curves are $14.5\mathring{A}/V$ for 1 tesla field and $43\mathring{A}/V$ for 3 tesla field.

kV/cm. The carrier concentration used in their calculations was $N = 5\times10^{17}$ $cm^{-3}$. Adjusted to that carrier concentration, the second order susceptibility for a 3-tesla magnetic field is about $215\mathring{A}/V$ compared to $240\mathring{A}/V$ of Ref. [9]. This shows that relatively weak magnetic fields in quantum wires can produce similar magnitudes of $\chi^{(2)}$ as rather strong electric fields in quantum wells.

Fig. 4 shows the dependence of the imaginary part of the first-order susceptibility $Im(\chi^{(1)})$ as a function of photon energy for two different values of the magnetic field. We have used the same wire dimensions and carrier concentrations for this plot as in the previous one. The same physics pertinent to the previous plot explains different broadening; however, the peaks are now attenuated because of averaging over different transition probabilities.

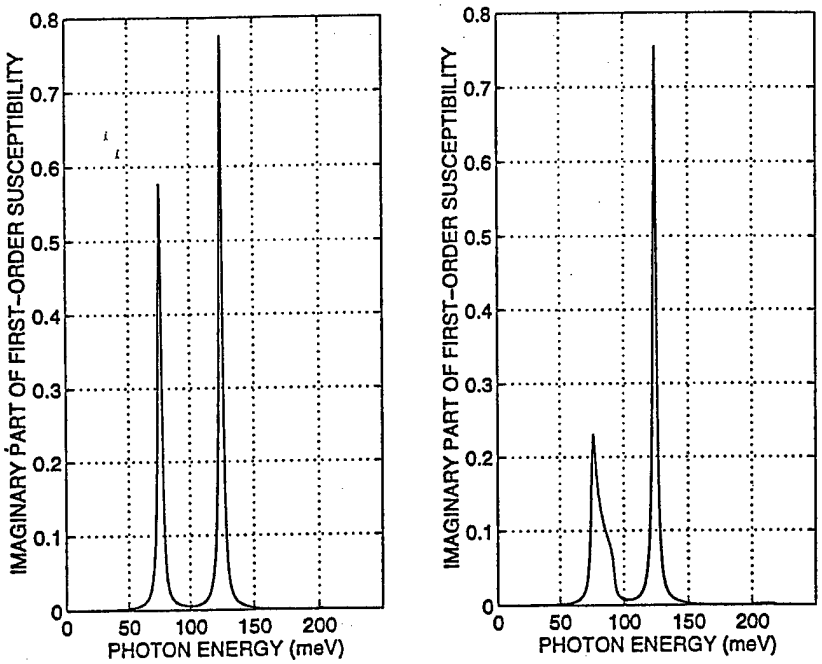



Figure 4: Imaginary part of the first-order susceptibility as a function of the photon energy for B=1 tesla (left panel) and B=3 tesla (right panel).

Since $Im(\chi^{(1)})$ is related to the absorption coefficient $\alpha(\omega)$ as

$$\alpha(\omega) = \frac{4\pi\omega}{c} Im(\chi^{(1)}), \tag{8}$$

one can estimate the absorption over the whole frequency range. At resonant photon energies of 72 meV and 124 meV, the absorption coefficient $\alpha = 1.5\ 10^4\ cm^{-1}$ and $\alpha = 4.5\ 10^4\ cm^{-1}$, respectively. It is clear from the figure, that the absorption coefficient at twice these frequencies, $\alpha(2\omega)$ is much less. This implies that a large portion of the pump energy at these resonant frequencies will be absorbed by the structure and converted into second harmonic signal which will not be significantly re-absorbed.

Another important factor for efficient second harmonic generation is *phase matching*. Since the refractive index $n(\omega)$ of most materials is frequency dependent, the following inequality holds $n(\omega) \neq n(2\omega)$. As a result, the coherence length $l_{coh} = \lambda_\omega/4(n_{2\omega} - n_\omega)$ for $GaAs$ (typical non-birefringent crystal)

varies between $10\mu$m and $100\mu$m. The efficiency of nonphase-matched interactions are about $10^{-5}$ times less than that of the phase-matched interactions over a length scale of 1 cm.

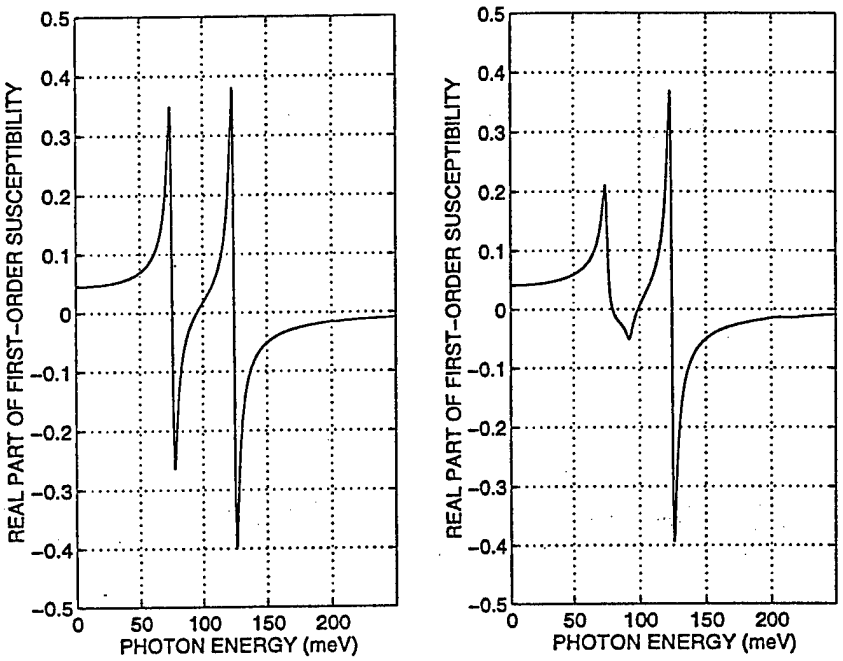



Figure 5: Real part of the first-order susceptibility as a function of the photon energy for B=1 tesla (left panel) and B=3 tesla (right panel).

Using a magnetic field as an additional degree of freedom, we may try to adjust $n(\omega)$. In Fig. 5 we present the dependence of the real part of the first-order susceptibility $Re(\chi^{(1)})$ as a function of photon energy for two different values of the magnetic field. Using the relation

$$\Delta n(\omega) = 2\pi Re(\chi^{(1)}), \tag{9}$$

and Eq. (8), one can determine the frequencies where two conditions simultaneously hold: $n(\omega) \approx n(2\omega)$ and $\alpha(\omega) >> \alpha(2\omega)$. For the 3-tesla field, this frequency $\hbar\omega$ corresponds to 75 meV.

## IV. Conclusion

We have theoretically studied second harmonic generation in a semiconductor quantum wire biased with a magnetic field. A strong second-harmonic component of the dielectric susceptibility, due to the dipoles associated with transitions between magneto-electric subbands, is found. We have also calculated absorption coefficient and refractive index in the appropriate frequency range to assess the efficiency of frequency conversion and insertion loss. We have shown that a magnetic field can be used as an additional degree of freedom in optimizing second harmonic generation efficiency. This may have important applications in nonlinear optics.

## Acknowledgement

This work is supported by the US Army Research Office under contracts DAAH04-95-1-0586 and DAAH04-95-1-0527. The authors are indebted to the Electrochemical Society, Inc. for partial travel support.

## REFERENCES

[1] see for example P.N. Butcher, D. Cotter, *The Elements of Nonlinear Optics,* (Cambridge University Press, 1990); A. Yariv, *Quantum Electronics,* (Wiley, N.Y., 1989); or Y.R. Shen, *The Principles of Nonlinear Optics,* (Wiley, N.Y., 1984).

[2]. E.J. Austin and M. Jaros, *Phys. Rev. B* **31**, 5569 (1985).

[3]. A. Svizhenko, A. Balandin and S. Bandyopadhyay, *J. Appl. Phys.*, (in press, 1997).

[4]. A. Balandin and S. Bandyopadhyay, *J. Appl. Phys.*, **77**, 5924 (1995); A. Balandin, Ph.D. dissertation, University of Notre Dame, 1996.

[5] This lack of symmetry is due to the transfer of a valency charge from one atom to the other bond-forming atom and resultant uneven charge distribution along the bond axis.

[6] *Handbook of Lasers,* Edited by R.J. Pressley (Chemical Rubber Co., Cleveland, 1971), p. 504.

[7]. Claude Weisbuch and Borge Vinter, *Quantum Semiconductor Structures: Fundamentals and Applications,* (Academic Press, Boston, 1991).

[8] S. Chaudhuri and S. Bandyopadhyay, *J. Appl. Phys.*, **71**, 3027 (1992).

[9]. M.M. Fejer, S.J.B. Yoo, R.L.Byer, A. Harwit, J.S. Harris, *Phys. Rev. Lett.*, **62**, 1041 (1989).

Presented at the Fourth International Symposium on Quantum Confinement: Nanoscale Materials and Devices, 191st Meeting of the Electrochemical Society, Montreal, Canada, July 4-9, 1997

# Confined Magneto-Polariton Propagation in Quantum Wires

A. Balandin and S. Bandyopadhyay
Department of Electrical Engineering
University of Nebraska
Lincoln, Nebraska 68588-0511 USA

**Abstract**

In this paper, we have calculated the refractive index of a quantum wire waveguide in the vicinity of an exciton-polariton resonance. The critical values of the exciton decay parameter, defining the onset of polariton transport regime, and the associated temperature were also found using the combination of a variational approach and a numerical solution. Our theoretical model allows us to include the effects of an external magnetic field. The results show that confinement of excitons to one dimension and the simultaneous application of a magnetic field may lead to the extension of the temperature and spatial limits of polariton transport. The magnetic field can be used to shift refractive index peaks in frequency thus providing a much-desired tuning capability.

## I. Introduction

Exciton dynamics in semiconductor quantum confined structures has always been attractive because of its potential applications in optoelectronics. One of the most interesting phenomenon related to excitons in such structures is the formation of exciton-polaritons. In the spectral region around an exciton resonance, a photon, absorbed by a semiconductor, linearly couples with an exciton to create a polariton. Exciton polaritons have been studied extensively in quantum wells by measuring photoluminescence and reflection,

as well as by picosecond time-of-flight measurements.[1,2] It was shown that polaritons are much more stable in quantum wells than in bulk,[3] and they are expected to be even more stable in quantum wires since the exciton binding energy and oscillator strength tend to increase with reducing dimensionality.[4]

The formation of polaritons modifies the transport of light through the medium. In particular, the medium of propagation becomes substantially more transparent[5,6] and the group velocity of light propagating along the waveguide approaches the speed of light in vacuum.[7] At the same time, there is a possibility of controlling polariton transport with an external field. In this paper, we argue that a relatively weak magnetic field is particularly attractive for tuning polariton transport because it increases exciton oscillator strength thus *extending the polariton regime* of energy transfer, while an electric field would decrease exciton oscillator strength and quench polariton transport.

Although existing theoretical models recognize modification of polariton transport due to spatial confinement, they account for it by using heuristically peaked values for exciton oscillator strength and binding energy. The authors are not aware of any attempts to include an external field, particularly magnetic, into consideration.

In this work, we present a model for calculating the exciton-polariton critical decay parameter and the refractive index of a quantum wire around a polariton resonance in the presence of a magnetic field. The decay parameter determines the regime of polariton transport. To our knowledge, this is the first study where the exciton longitudinal-transverse (LT) splitting and exciton resonance frequency - which define the polariton dispersion - are found in a non-ad-hoc manner using the combination of a variational approach and an exact numerical solution of the Schrödinger equation. The calculations are performed for a quantum wire with finite lateral dimensions subjected to a magnetic field.

The rest of the paper is organized as follows. In section II, we establish the polariton dispersion relation used throughout the model; section III presents

the results of calculation of the LT splitting and oscillator strength of the exciton transition in a quantum wire subjected to a magnetic field; in section IV, we examine the exciton critical damping for the onset of the polariton transport regime and calculate refractive index of the wire in the vicinity of exciton resonance. Conclusions are given in section V of the paper.

## II. Polariton dispersion

We consider an array of parallel $GaAs$ quantum wires of rectangular cross section separated by infinite potential barriers so that wavefunctions of the excitons from different wires do not overlap. In such a structure, excitons are free to move along the wire axes but are confined in perpendicular directions. The lateral dimension of each wire is comparable to the exciton Bohr radius. The cladding material is assumed to have a similar refractive index so that we can ignore image charges of the exciton and associated dielectric confinement effects. Under these conditions, the dispersion relation of the exciton polaritons can be determined for each separate wire. The multiple wire structure in this case merely forms a waveguide structure analogous to that considered in Ref. [8].

Most theoretical models for exciton polaritons (both in bulk material and nanostructures) embody a semi-classical approach and utilize the dispersion relation of a polariton derived for a single electric-dipole-active exciton resonance.[5,8] Here, we adopt the same philosophy and consider electromagnetic waves propagating through an array of quantum wires with a wavevector $k$ parallel to the wire axis. This choice of the direction of propagation allows for a spatial dispersion of the light waves. In the opposite case of light propagating normal to the wire axis, the translational motion of excitons is suppressed and the spatial dispersion effects do not occur.

In the long-wave approximation ($kL_{y,z} < 1$, where $L_{y,z}$ are the wire lateral dimensions) the array interacts with light waves like an effective medium, and the dielectric function in the vicinity of an isolated exciton resonance can be

written as[9]

$$\epsilon(\omega,k)=\epsilon_o+\frac{2\epsilon_o\omega_{LT}\omega_o}{\omega_o^2-\omega^2+\hbar k^2\omega_o/M-i\omega\gamma}. \qquad (1)$$

where $\epsilon_o$ is the background dielectric constant (contribution made by other resonances), $\omega$ is the frequency of light, $\omega_{LT}$ is the longitudinal-transverse splitting of the exciton related to its oscillator strength, $\omega_o$ is the exciton resonant frequency at $k=0$, $M=m_e+m_h$ is the translational mass of an exciton, and $\Gamma\equiv\hbar\gamma$ is the exciton damping parameter. Here we have assumed parabolic wavevector dependence of the exciton frequency $\hbar\omega_t(k)=\hbar\omega_o+\hbar^2k^2/2M$, with the caveat that this is valid only in weak magnetic fields when the magnetic length $l_m$ $(=\sqrt{\hbar/eB})$ is much larger than the transverse dimensions of the wire. In the formula above, the quantities $\omega_{LT}\equiv\omega_{LT}(L_{y,z},B)$ and $\omega_o\equiv\omega_o(L_{y,z},B)$ are the functions of the wire lateral dimensions and a magnetic field. The exciton damping constant is considered to be independent of the magnetic field since it is known that *energy-averaged* phonon-interaction rates in quantum wires are not terribly sensitive to a magnetic field. In any case, Eq. (1) is a good approximation when the magnetic field applied to the system is relatively weak: $l_m>L_{y,z}$. This equation relates $\omega$ and $k$ and is the sought-after dispersion relation of a polariton.

Before we can go further into polariton transport properties, we have to calculate $\omega_{LT}$ and $\omega_o$ as the functions of wire dimensions $L_{y,z}$ and a magnetic flux density B. This is discussed in the next section.

## III. Longitudinal-transverse splitting

Let us assume that the infinite potential barriers of the quantum wire are located at $y=\pm L_y/2$ and $z=\pm L_z/2$. A magnetic field is applied along the $z$-direction (see inset to Fig. 1). To simplify the calculations, we assume strong quantum confinement of the carriers which enables us to factorize an exciton wavefunction into the product of electron and hole wave functions. Moreover, we limit our consideration to systems with relatively large dielec-

tric constants so that all Coulomb interactions are strongly screened. This assumption, together with the hard-wall boundary condition, allows a coordinate separation. Consequently, the wave function of an exciton in the vicinity of subband bottom (with center-of-mass momentum $\hat{P_X} \approx 0$) is given by[4]

$$\begin{aligned}\Psi \equiv \Psi(x, y_e, y_h, z_e, z_h) &= g_t(x,\eta)\psi_e(y_e, z_e)\psi_h(y_h, z_h) \\ &= g_t(x,\eta)\phi_e(y_e)\phi_h(y_h)\chi_e(z_e)\chi_h(z_h),\end{aligned} \tag{2}$$

where $g_t(x,\eta)$ is chosen to be the Gaussian-type "orbital" function:

$$g_t(x,\eta) = \frac{1}{\eta^{1/2}}(\frac{2}{\pi})^{1/4} e^{-(x/\eta)^2} \tag{3}$$

in which $\eta$ is a variational parameter which defines the exciton size ("longitudinal length"), and $x$ is the relative electron-hole coordinate. The subscripts in $x_{e,h}, y_{e,h}, z_{e,h}$ identify them as electron or hole coordinates. The variables $\chi_{e,h}(z_{e,h})$ are the z-components of the wave functions which are not affected by the magnetic field. They are given by particle-in-a-box states. The electron and hole wave functions along the $y$ direction, $\phi_{e,h}(y_{e,h})$, are to be calculated numerically when a magnetic field is present. This is done by solving the one-particle Schrödinger equation using a finite difference method.[10]

In order to find an exciton "length" $\eta$, we use the variational approach of minimizing the energy given by $< \Psi|\hat{H^X}|\Psi >$, where the exciton Hamiltonian is

$$\begin{aligned}\hat{H^X} &= \frac{\hat{P_X}^2}{2M} + \frac{\hat{p_x}^2}{2\mu} + \frac{\hat{p_{y_e}}^2 + \hat{p_{z_e}}^2}{2m_e} + \frac{\hat{p_{y_h}}^2 + \hat{p_{z_h}}^2}{2m_h} \\ &+ \frac{eB(y_e - y_h)}{M}\hat{P_X} + eB(y_e/m_e + y_h/m_h)\hat{p_x} + \frac{e^2B^2}{2}(y_e^2/m_e + y_h^2/m_h) \\ &+ U_C(x_e, x_h, y_e, y_h, z_e, z_h) + U_S(y_e, y_h, z_e, z_h).\end{aligned} \tag{4}$$

Here we have chosen the Landau gauge $\vec{A} = (-By, 0, 0)$. The quantities $m_e, m_h$, are the effective masses of electrons and holes respectively, $1/\mu (= 1/m_e + 1/m_h)$ is the exciton's reduced mass, $U_C(x_e, x_h, y_e, y_h, z_e, z_h)$ is the electron-hole Coulomb interaction term, $U_S(y_e, y_h, z_e, z_h)$ is the spatial confinement potentials for electrons and holes along $y$ and $z$ directions.

Details of the variational procedure, calculations of $\eta$, etc., can be found in some of our earlier work.[4]

We can find the oscillator strength of the exciton transition $\alpha_o$ and the LT splitting by evaluating the momentum matrix element which is given as

$$|M_{cv}^X|^2 = |\frac{1}{2\pi}\int dk g_t(x,\eta) M_{cv}(k)|^2, \tag{5}$$

where $M_{cv}$ is the valence-band to conduction-band dipole matrix element. and $k$ is again the wave vector along the unconfined direction of the wire. When the $k$ dependence of $M_{cv}$ is neglected, Eq. (5) reduces to the simple expression

$$|M_{cv}^X|^2 = |M_{cv}|^2 |g_t(x=0,\eta)|^2. \tag{6}$$

The exciton oscillator strength per unit length can be written as follows

$$\alpha_o = \frac{2}{m_o \hbar\omega_o}|M_{cv}^X|^2. \tag{7}$$

Here $\hbar\omega_o = E_G + E_{e1} + E_{hh1} - min < \Psi|\hat{H}|\Psi >$ is the exciton ground state energy, $E_G$ is the fundamental bandgap of the bulk material, $E_{e1}$, $E_{hh1}$ are the lowest electron and the highest heavy hole magneto-electric subband bottom energies in a quantum wire measured from the bottom of the bulk conduction band and the top of the bulk valence band, and $m_o$ is the free electron mass. The exciton LT splitting $\Omega_{LT} \equiv \hbar\omega_{LT}$ can now be written as

$$\hbar\omega_{LT} = \frac{2\pi\alpha_o\hbar\omega_o}{\epsilon_o} = \frac{4\pi}{m_o\epsilon_o}|M_{cv}^X|^2. \tag{8}$$

In Fig. 1 we present the LT splitting calculated for different wire dimensions and magnetic flux densities. The physical parameters used for the calculations correspond to a $GaAs$ quantum wire with $\epsilon = 12.9\epsilon_o$, $E_G = 1.515$eV, $m_e = 0.067m_o$, $m_h = 0.5m_o$, where $m_o$ is free electron mass and $\epsilon_o$ is electrical permitivitty of free space, $E_{e1}$, $E_{hh1}$ are calculated numerically following the prescriptions of Refs. [10, 11]. One can see from the figure, that the exciton splitting is sensitive to the spatial confinement and increases by about 60 % when the wire width decreases from 500 Å to 50 Å. A magnetic field

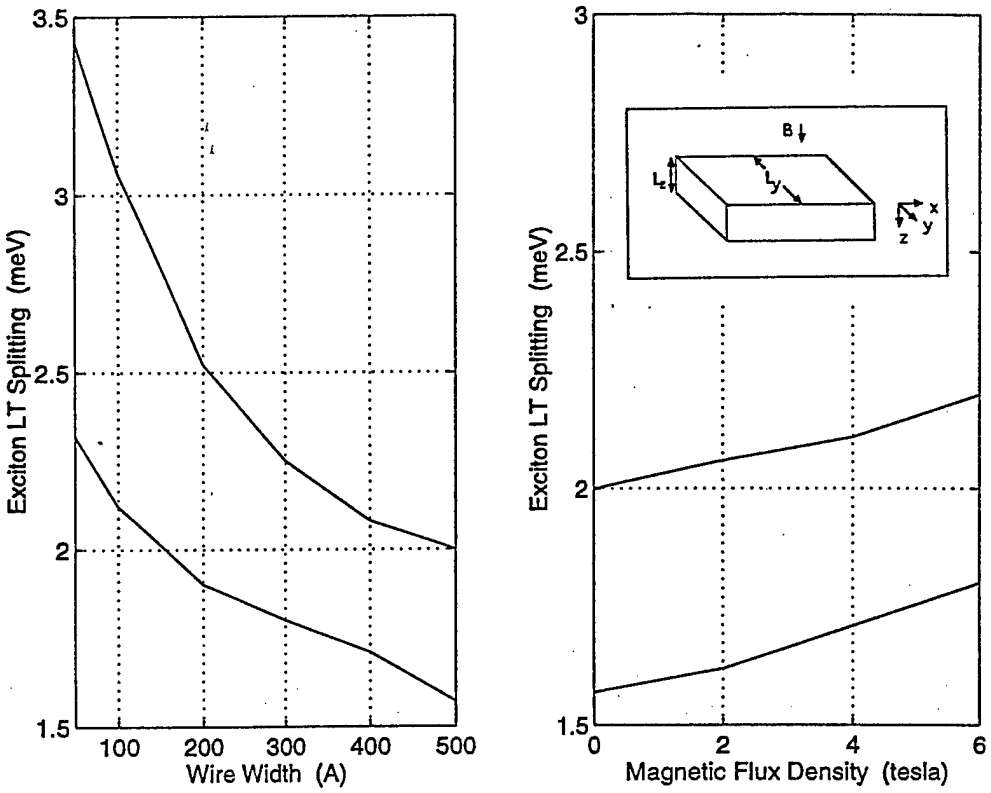



Figure 1: Longitudinal-transverse splitting of an exciton as a function of wire width (left pannel); and as a function of magnetic flux density for a 500 Å wide wire (right panel). For both panels, the upper and lower curves correspond to 200 Å and 500Å thick wires, respectively.

also increases the splitting (and oscillator strength) thus making the exciton polaritons more stable.

## IV. Refractive index and decay parameter

The experimentally observed higher transparency of the medium of propagation in polariton transport regime has been attributed to certain features of the dispersion law for excitonic polaritons and to the fact that polariton transport by itself cannot cause true absorption. In order for absorption to occur, polaritons have to be scattered inelastically, e.g., by phonons. The onset of polariton transport through some structure is governed by the exciton

polariton coherence length related to the exciton decay parameter $\Gamma \equiv \hbar\gamma$. It has been shown, both experimentally and theoretically, that there exists a critical value of the exciton decay parameter, $\Gamma_c$, which corresponds to a change in the nature of absorption.[5,6] Here we intend to examine the influence of spatial confinement and a magnetic field on this parameter and calculate the refractive index of a quantum wire in the polariton regime.

Confining ourselves to the TE waves relevant to light propagation in the medium, we can write polariton dispersion in the following form

$$\epsilon(\omega, k) = \frac{c^2k^2}{\omega^2} \equiv n^2. \tag{9}$$

Combining this equation with Eq. (1), and after some algebra, we obtain

$$\frac{\hbar\omega_o\omega^2}{Mc^2}n^4 + (\omega_o^2 - \omega^2 - i\gamma\omega - \epsilon_o\frac{\hbar\omega_o\omega^2}{Mc^2})n^2 - \epsilon_o(\omega_o^2 - \omega^2 - i\gamma\omega + 2\omega_{LT}\omega_o) = 0. \tag{10}$$

This equation can be solved for two sets of the refractive indices, $n_1$ and $n_2$, corresponding to different transverse polariton branches. It also follows from Eq. (10) that if the damping parameter $\Gamma$ becomes larger than the critical value

$$\Gamma_c \equiv \Gamma_c(L_{y,z}, B) = 2\hbar\omega_o\sqrt{\frac{2\epsilon_o\hbar\omega_{LT}}{Mc^2}}, \tag{11}$$

then only one light wave mode can propagate in the medium, since there is only one real solution for $n$. This is the boundary of the polariton propagation regime. The critical value comes about because of the term $\hbar k^2\omega_o/M$ associated with spatial dispersion effects. The physical importance of the critical damping can be illustrated by the following example. It was shown[6] that when the damping exceeds the critical value, the integral absorption is independent of $\Gamma$ and proportional to the oscillator strength of transitions (non-polariton regime). When $\Gamma < \Gamma_c$, the integral absorption depends on $\Gamma$ linearly, and decreases with decreasing damping.

Using the results from the previous section, we can calculate $\Gamma_c$ for different values of wire widths and magnetic field. In Fig. 2 we present the critical exciton damping (decay) parameter as a function of wire width. It is

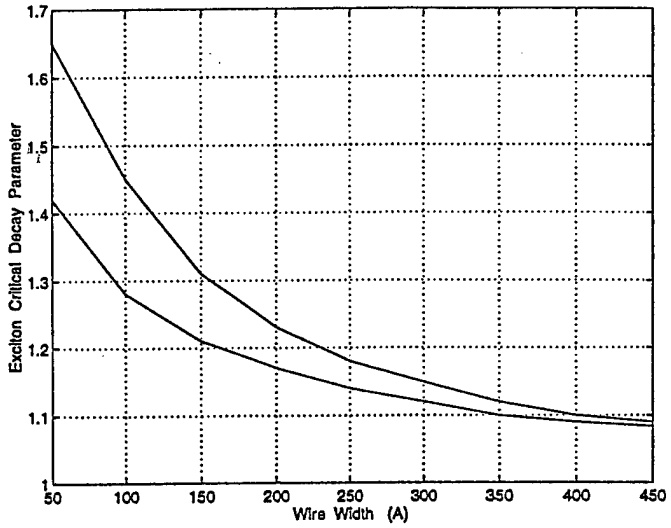



Figure 2: The critical exciton decay parameter (damping) as a function of wire width. The upper and lower curves correspond to the thickness along the z-direction of 200 Å and 300 Å, respectively.

normalized by the value of the decay parameter of a very wide wire ($L_y$=700 Å) which is wide enough to be approaching the 2D limit. For this wide wire, $\Gamma_c \approx 1.01$ meV. Combining Eqs. (7, 8, 11) we can also estimate the magnetic field dependence of the critical decay parameter using the formula $\Gamma_c(B)/\Gamma_c(0) = (\omega_o(B)/\omega_o(0))\sqrt{\omega_{LT}(B)/\omega_{LT}(0)}$. Although not shown here, the magnetic field dependence of the critical parameter is weak; it increases only 5% at a magnetic flux density of 5 tesla. The strong dependence of the critical damping on the wire width may lead to a pronounced modification of the integral (total) absorption of systems consisting of a number of narrow quantum wires.

We can now find a temperature which corresponds to the critical damping from the relation $\Gamma_c = \Gamma_o + \Gamma_{ph}(T_c)$, where $\Gamma_o$ is the damping associated with the impurity and other temperature-independent elastic scattering, while $\Gamma_{ph}(T_c)$ represents interactions with acoustic and optical phonons. By increasing $\Gamma_c$ one can increase $T_c$ which defines the onset of polariton transport and, as a consequence, higher transparency.

Since there is no data available on quantum wires, we assume that the half-width at half maximum (HWHM) of the exciton resonance in a quantum wire is the same as in a 200 Å thick $GaAs/AlGaAs$ quantum well. Using the approximation of Ref [11] we may write for our case (energy units are meV)

$$\Gamma_c = \Gamma_o^+ + 0.00147T_c + 4.0(e^{\hbar\omega_{ph}/k_BT_c} - 1)^{-1} + \Gamma_{imp}e^{-E_b/k_BT_c}, \qquad (12)$$

where $\hbar\omega_{ph} = 36$ meV is a longitudinal optical phonon energy, $k_B$ is the Boltzman constant, $E_b \approx 10$ meV is the everage binding energy for donor impurities in GaAs, $\Gamma_{imp} = 0.75$ meV is a linewidth due to fully ionized impurity scattering, $\Gamma_o^+ \approx 0.45$ meV is the linewidth due to inhomogeneous fluctuations of the wire thickness. The values chosen for the various parameters are typical of experimental systems reported in the literature.

**Table I. Critical temperature vs. wire width**

| $L_y, (\AA)$ | 500 | 300 | 100 |
|---|---|---|---|
| $\Gamma_c, (meV)$ | 1.09 | 1.16 | 1.45 |
| $T_c, (K)$ | 118 | 138 | 169 |

Solving Eq. (12) for the temperature $T_c$ for each value of $\Gamma_c(L_{y,z}, B)$, we are able to obtain the dependence of the critical temperature on the wire width. The thickness of the wire, $L_z$, was fixed at 200Å for this calculation. As one can see from Table I, that the critical temperature $T_c$, that defines the onset of polariton transport, can be controlled over a wide range by changing the wire width $L_y$.

Now let us assume that $\Gamma < \Gamma_c$ (exciton polariton regime) and find the refractive index of the quantum wire in the vicinity of polariton resonance. In order to do this, we make use of Pekar's additional boundary condition (total polarization is zero at the boundary) and write the effective refractive index as

$$n_{eff} = \frac{n_1 n_2 + \epsilon_o}{n_1 + n_2}. \qquad (13)$$

In Fig. 3, we present the real (upper panel) and imaginary (lower panel) parts of the refractive index of a quantum wire with thickness 200 Å and

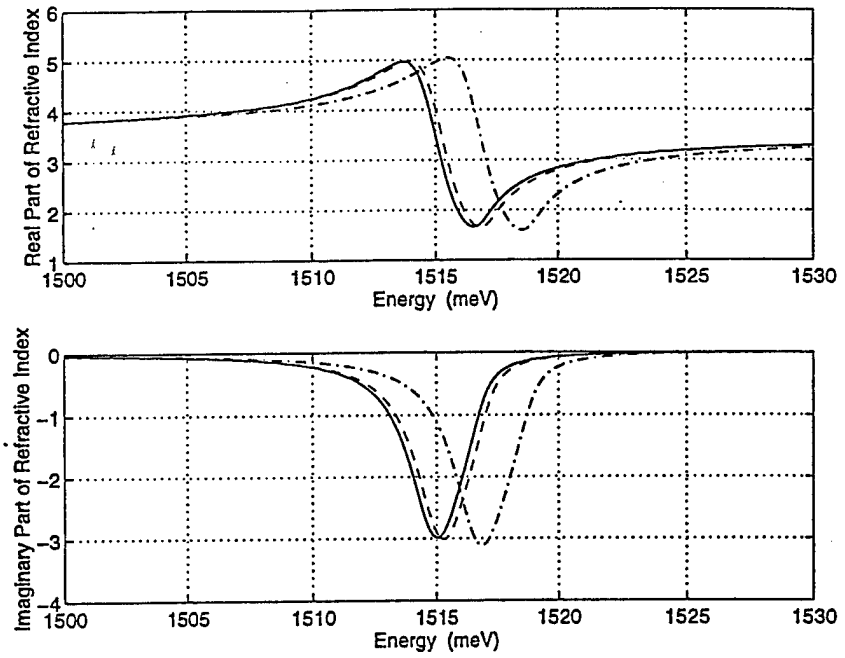



Figure 3: The refractive index of the wire in the vicinity of exciton resonance. The solid, dashed and dash-dotted curves correspond to a 0, 1, and 5 tesla magnetic flux density, respectively.

width 300 Å. An external transverse magnetic field is applied along the thickness. The decay parameter is chosen to be $\Gamma = 1\text{meV}$ which is less than $\Gamma_c$ for the given wire dimensions. As one can see from the figure, the maximum of the real part of the refractive index is as large as 4.97 at zero field and 5.04 at 5 tesla magnetic field. It is about 1.4 times larger than that of the bulk material. The refractive index attains its maximum value at a photon energy slightly lower than the resonance energy $\hbar\omega_o(L_{y,z}, B)$ at any given magnetic field B. The minimum value of the refractive index, which is 1.66 for zero field and 1.59 for 5 tesla, is located at a frequency of $\omega_o(L_{y,z}, B) + \omega_{LT}(L_{y,z}, B)$. It is interesting to note that efficient waveguiding can be achieved in the spectral range where the real part of the refractive index increases. However, in this region, the imaginary part of the refractive index (extinction coefficient) also peaks and this increases the transmission

loss. The mitigating factor in all this is that the two peaks do not occur at exactly the same frequency so that an optimal region for optical waveguiding exists.

A magnetic field of 5 tesla blue-shifts the refractive index peak by 2 meV. According to Ref. [13], approximately the same magnitude of an opposite red-shift can be achieved by applying an electric field of about $4\times10^4$ V/cm. However, the electric field leads to a 15-20 % increase in exciton radius and a concomitant decrease in the binding energy. This, in turn, causes a decrease in exciton LT splitting and makes exciton polaritons less stable. Note that while an electric field will tend to ionize an exciton by pulling the electron and hole apart, a magnetic field has the opposite effect. It squeezes the electron and hole even tighter together and increases the binding energy. Therefore, the magnetic field can be used to advantage in this context since it shifts the peaks in frequency while actually increasing polariton stability. This frequency tuning capability, acquired without a penalty in polariton stability, is obviously very attractive and has device applications.

## VI. Conclusions

In this paper, we have calculated the refractive index of a quantum wire waveguide in the vicinity of polariton resonance. The critical values of the exciton decay parameter and associated temperature were also found taking into account the effects of spatial confinement and an external magnetic field. Our results show that confinement of excitons to one dimension and the application of a magnetic field may lead to the extension of temperature and spatial limits of polariton transport. The magnetic field can be used to shift refractive index peaks in frequency - without compromising polariton stability - thus providing a much-desired tuning capability.

**Acknowledgement**: This work was supported by the US Army Research Office under grant number DAAH04-95-1-0527. The authors are indebted to the Electrochemical Society, Inc. for providing travel support.

# REFERENCES


[1]. E.L. Ivchenko, V.P. Kochereshko, P.S. Kopev, V.A. Kosobukin, I.N. Uraltsev, D.R. Yakovlev, *Solid State Commun.*, **70**, 529 (1989).

[2]. K. Ogawa, T. Katsuyama, H. Nakamura, *Phys. Rev. Lett.*, **64**, 796 (1990).

[3]. T. Katsuyama, K. Ogawa, *Semicond. Sci. Technol.*, **5**, 446 (1990).

[4]. A. Balandin and S. Bandyopadhyay, *Phys. Rev. B*, **52**, 8312 (1995); A. Balandin and S. Bandyopadhyay, *Phys. Rev. B*, **54**, 5712 (1996).

[5]. V.A. Kosobukin, R.P. Seisyan, S.A. Vaganov, *Semicond. Sci. Technol.*, **8**, 1235 (1993).

[6]. N. Akhmediev, *Zh. Eksp. Teor. Fiz.*, **79**, 1554 (1980).

[7]. K. Oimatsu, T. Iida, S. Nishimura, K. Ogawa, T. Katsuyama *Journal of Luminescence*, 48, 713 (1991).

[8]. T. Katsuyama, S. Nishimura, K. Ogawa, T. Sato, *Semicond. Sci. Technol.*, **8**, 1226 (1993).

[9]. see for example E.L. Ivchenko in *Excitons* (North-Holland, Amsterdam, 1992), edited by E.I. Rashba and M.D. Sturge, p.141.

[10]. S. Chaudhuri, S. Bandyopadhyay, *J. Appl. Phys.*, **71**, 3027 (1992).

[11]. A. Balandin and S. Bandyopadhyay, *J. Appl. Phys.*, **77**, 5924 (1995).

[12]. J. Lee, E.S. Koteles, M.O. Vassell, *Phys. Rev. B*, **33**, 5512 (1986).

[13]. D.A.B. Miller, D.S. Chemla, T.C. Damen, A.C. Gossard, W. Wiegmann, T.H. Wood, C.A. Burrus, *Phys. Rev. Lett.*, **53**, 2173 (1984); *Phys. Rev. B*, **32**, 1043 (1985).

# Theoretical studies of the effects of a magnetic field on excitonic nonlinear optical properties of quantum wires

A. Balandin* and S. Bandyopadhyay*
*Department of Electrical Engineering, University of Nebraska, Lincoln, Nebraska 68588*
(Received 12 February 1996)

The magnetic-field dependence of the third-order excitonic nonlinear susceptibility $\chi^{(3)}$ in a quantum wire is explored within the rotating wave approximation. Both the real and imaginary parts of $\chi^{(3)}$, arising from population saturation of the excitonic state under optical pumping, are calculated for a GaAs wire as a function of magnetic field and pump-probe detuning frequencies. The imaginary part of $\chi^{(3)}$ exhibits a negative peak associated with the bleaching of the excitonic resonance and a positive, broad, off-resonance absorption peak associated with biexciton formation. The amplitude, line shape, and spectral frequency of both these peaks can be modulated by a magnetic field which indicates the possibility of using such a field to probe the mechanism underlying optical nonlinearity in a quantum wire. Furthermore, the field can also be used to tune the optical nonlinearity over a range of frequencies which has device applications. [S0163-1829(96)04228-2]

## I. INTRODUCTION

It is well known that quasi-one-dimensional systems (quantum wires) exhibit giant third-order nonlinear susceptibility $\chi^{(3)}$ under optical pumping because of quantum confinement of excitons and polyexcitonic complexes.[1,2] A magnetic field further increases the confinement by localizing the electron and hole wave functions, leading to even larger $\chi^{(3)}$. This allows one to modulate the nonlinear refractive index and absorption (or gain) in quantum wires with a magnetic field, thereby opening up the possibility of realizing *externally tunable* couplers, limiters, phase shifters, switches, etc. Furthermore, the field can also be used as an experimental tool to extract the specific mechanism responsible for the optical nonlinearity in the system.

In this paper, we will investigate the effects of a magnetic field on optical nonlinearity (and the associated $\chi^{(3)}$) in a quantum wire caused by exciton-exciton interaction and formation of excitonic molecules (specifically biexcitons). This interaction is likely to be the dominant mechanism for optical nonlinearity, and the leading contribution to $\chi^{(3)}$ in quantum wires of most technologically important semiconductors. Recently some researchers[2,3] reported experimental observations of giant optical nonlinearity in quantum wires which they attributed to this mechanism. The enhanced nonlinearity is undoubtedly caused by quantum confinement which increases the binding energy of all excitonic complexes. Additionally, the oscillator strength for the lowest-energy exciton-to-biexciton transition increases and gives rise to huge third-order susceptibilities. This oscillator strength is already significant because the biexciton radius is very large,[4] and a second photon (two-photon absorption) in the volume of an excitonic molecule can be easily found and absorbed to create the molecule.

In Sec. II, we outline the theoretical framework that was used for calculating $\chi^{(3)}$ associated with population saturation of the excitonic state and biexciton formation. Section III presents the results of our numerical computation followed by a discussion. We also compare the results that we obtain (in the absence of any magnetic field) with the theoretical calculations of Ref. 4, and with available experimental data. Conclusions are given in Sec. IV.

## II. THEORETICAL MODEL

We consider a rectangular GaAs quantum wire of the geometry shown in the inset of Fig. 1. An external magnetic field of flux density $B$ is applied perpendicular to the wire axis, as indicated in the inset. In order to calculate $\chi^{(3)}$ for this system, we make the following assumptions: (i) $\chi^{(3)}$ is measured in a nondegenerate pump-and-probe spectroscopy experiment with nearly resonant pumping of the excitonic state; (ii) the exciton gas is sufficiently dilute that higher-order complexes (beyond the biexcitonic state) can be neglected; (iii) the rotating wave approximation[5] is valid; and (iv) in the frequency range of interest, the lowest-lying states are the major contributors to $\chi^{(3)}$ and therefore we can treat the system approximately as a two-level system.

Following Ishihara[2] and Madarasz *et al.*,[4] we can write the third-order susceptibility as follows:

$$\chi^{(3)}=\frac{-2}{\pi\sqrt{2\pi}}\frac{\tau}{\eta^2}\frac{N_0e^4}{m_0^2\omega_{g0}^4}E_p^2\left[\frac{1}{(\omega_1-\omega_{g0}+i\Gamma_{g0})}-\frac{1}{(\omega_1-\omega_{g0}+\omega_b+i\Gamma_{bg})}\right]\times\sum_{r=1}^{2}\left\{\frac{1}{\hbar^3(\omega_r-\omega_2+i\gamma)}\left[\frac{1}{(\omega_{g0}-\omega_2+i\Gamma_{g0})}+\frac{1}{(\omega_r-\omega_{g0}+i\Gamma_{g0})}\right]\right\}, \tag{1}$$

where $\omega_2$ and $\omega_1$ are the pump and probe frequencies, $\hbar\omega_{g0}$ is the exciton ground-state energy, $\hbar\omega_b$ is the biexciton binding energy, $m_0$ is the rest mass of a free electron, and $N_0$ is the average areal density of unit cells. The quantities $\Gamma_{ij}$ and $\gamma$ are the transverse and longitudinal broadening parameters (or damping constants), and $E_p$ is the Kane matrix

 © 1996 The American Physical Society

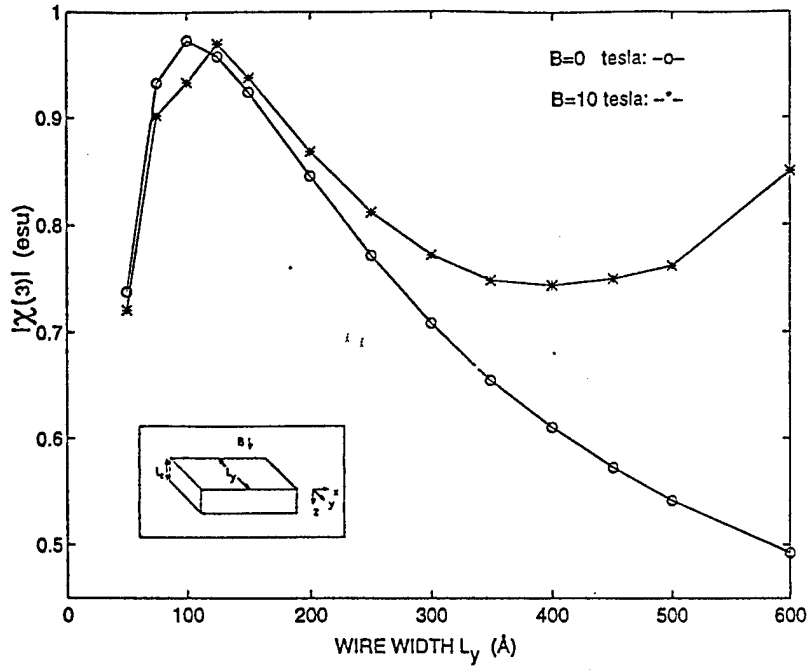



FIG. 1. The absolute value of the third-order susceptibility $\chi^{(3)}$ (for resonant excitation) as a function of wire width for two different values of magnetic field. The inset shows the wire geometry and the orientation of the magnetic field. The wire thickness along the $z$ direction is 200 Å.

element. The indices $i$ or $j$ indicates system ground state (0), exciton ground state ($g$), and biexciton ground state ($b$). The values of the parameters used are listed in Table I, and correspond to GaAs. Parameters $\eta$ and $\tau$ physically correspond to the exciton and biexciton correlation lengths (electron hole and hole hole mean separations in the two cases), and have to be determined variationally for each magnetic-field strength and for each set of wire dimensions following the prescription given in Refs. 6 and 7.

The exciton ground-state energy $\hbar\omega_{g0}$ is defined as follows:

$$E_g^X=\hbar\omega_{g0}=E_G+E_{e1}+E_{\mathrm{hh1}}-E_B^X, \tag{2}$$

where $E_G$ is a bulk band gap of the material, $E_{e1}$ and $E_{\mathrm{hh\,1}}$ are the lowest electron and the highest heavy-hole magneto-electric subband bottom energies in a quantum wire (measured from the bottom of the bulk conduction band and the top of the bulk valence band) respectively, and $E_B^X$ is the ground-state exciton binding energy which is also determined variationally.[6,7]

One should note from Eq. (1) that $\chi^{(3)}$ is a strong function of the transverse and longitudinal broadening parameters $\Gamma_{ij}$ and $\gamma$. Physically, $\gamma$ is related to the population decay rate of the excitonic states. The smaller the value of $\gamma$, the larger the lifetime of excitons and the higher the probability of forming a biexciton in a two-step photon absorption. The transverse broadening parameters $\Gamma_{ij}$ represent, for $i\neq j$, the phenomenological coherence decay rate of the $ij$ transition, while, for $i=j$, they describe the population decay of the state $i$. The population decay rate, in its turn, is determined by the dominant scattering mechanism in the sample. In most cases, the values of $\Gamma_{ij}$ and $\gamma$ are difficult to obtain experimentally, and fairly difficult to estimate theoretically. Moreover, these parameters could be strong functions of the confinement, population density of excitons, magnetic field, and temperature. In view of little experimental data available, and in order to simplify the calculations, we assume that $\Gamma_{ij}=\Gamma$ for all $i$ and $j$.

TABLE I. Physical parameters for GaAs.

| |
|---|
| $E_G=1.519$ eV |
| $\hbar\Gamma=3$ meV |
| $E_p=23$ eV |
| $N_0=7.89\times10^{14}/\mathrm{cm}^2$ |

Since in this work we are interested in the modulation of the nonlinear response of quantum wires with a magnetic field, the influence of the field on the above parameters is especially important. The value of $\Gamma$ in quantum wires is primarily determined by carrier-phonon interaction.[8] As shown in Ref. 8, the scattering rates associated with these interactions can be affected by a magnetic field at any given kinetic energy of an electron or hole. However, when the rates are averaged over the energy, the magnetic-field dependence turns out to be quite weak. As a first approximation, we can therefore consider the rates to be independent of the magnetic field. We also neglect thermal broadening of the damping parameters, since it is less important in quantum confined systems than in bulk.[9] An important property of Eq. (1) is the following. If all the transverse relaxation parameters are assumed to be equal (like in our case) and the biexciton binding energy ($\hbar\omega_b$) approaches zero, then $\chi^{(3)}$ vanishes. This is a reflection of the well-known fact that noninteracting ideal independent bosons do not show any nonlinearity.[10] Consequently, exciton-exciton interaction, leading to biexciton formation, is necessary for the existence of nonlinearity.

A calculation of the excitonic contribution to $\chi^{(3)}$ requires that the exciton and biexciton binding energies be obtained first. Additionally, the parameters $\eta$ and $\tau$ need to be found. For details of computing these energies and these parameters in the case of a quantum wire subjected to a magnetic field, we refer the reader to our past work.[6,7,11] Once these quan-

tities are evaluated, we can calculate $\chi^{(3)}$ from Eq. (1) as a function of magnetic field, wire width, and pump-probe detuning frequencies.

## III. RESULTS AND DISCUSSION

All results in this paper are pertinent to GaAs quantum wires. In Fig. 1, we plot the absolute value of $\chi^{(3)}$ as a function of the wire width for a fixed wire thickness of 200 Å , with and without a magnetic field. The susceptibility peaks at about 0.97 esu, corresponding to a wire width of about 100–120 Å . The sharp drop at smaller wire widths is caused by a fast rise in electron and hole confinement energies $E_{e1}$ and $E_{hh1}$ with shrinking wire width. This rise is faster than the rise in the exciton binding energy $E_B^X$, which eventually leads to a decrease in $|\chi^{(3)}|$. For dimensions larger than 120–150 Å , the ground-state energy varies little, and the behavior of $\chi^{(3)}$ is primarily determined by the effective exciton and biexciton correlation lengths $\tau$ and $\eta$. When no magnetic field is present, both $\tau$ and $\eta$ increase with increasing wire width, but $\eta$ increases at a faster rate. Consequently, the term $(\tau/\eta^2)$ decreases monotonically with increasing wire width, making $\chi^{(3)}$ decrease. This decrease is somewhat offset by the variation of the ground-state exciton energy, which causes the roll-off rate to be more gentle than at small wire widths. For a nonzero magnetic field (of 10 T), the exciton radius $\eta$ in a quantum wire has a *nonmonotonic* dependence on wire width which results in a well-resolved maximum in $\eta$. This rather surprising behavior was reported by us earlier,[11] and explained in terms of the complementary roles of electrostatic and magnetostatic confinement. The biexciton radius $\tau$ also has a maximum, but it is much broader and shallower than the one associated with $\eta$. Consequently, there exists a minimum in the ratio $\tau/\eta^2$ which causes $\chi^{(3)}$ to exhibit a nonmonotonic dependence on the wire width (past the maximum) when a magnetic field is present. This accounts for the broad valley in the curve when a magnetic field of 10 T is applied.

In Figs. 2–4 we have calculated $\mathrm{Im}\chi^{(3)}$ for a two-beam experiment in which the frequency of one beam, the pump, is fixed, and that of the other, the probe, is allowed to vary over a frequency range of $\hbar\Delta\omega=40$ meV centered around the pump frequency. In Figs. 2 and 4, the pump frequency is chosen to be resonant with the exciton ground-state transition, and in Fig. 3 the pump is detuned from the exciton resonance by a frequency $-\Gamma/(2^{1/2}\hbar)$. It is important to remember that since the ground-state exciton binding energy is a function of magnetic field, the pump should be retuned every time the magnetic field changes. In all figures, the imaginary part of the third-order susceptibility is plotted for four values of magnetic field. The significance of $\mathrm{Im}\chi^{(3)}$ is in that it is proportional to the differential change in the optical transmission or in the absorption coefficient $\Delta\alpha$. This relation is given by the formula[12]

$$\mathrm{Im}\chi^{(3)}=\frac{c^2 n_0^2 \Delta\alpha(\omega)}{8\pi^2\omega I(\omega)}, \tag{3}$$

where $n_0$ is a linear refractive index, $c$ is the speed of light, and $I(\omega)$ is the intensity of a resonant monochromatic light beam. Positive peaks in $\mathrm{Im}\chi^{(3)}$ will correspond to strong absorption and negative peaks to the spectral regions of strong transmission (bleaching bands).

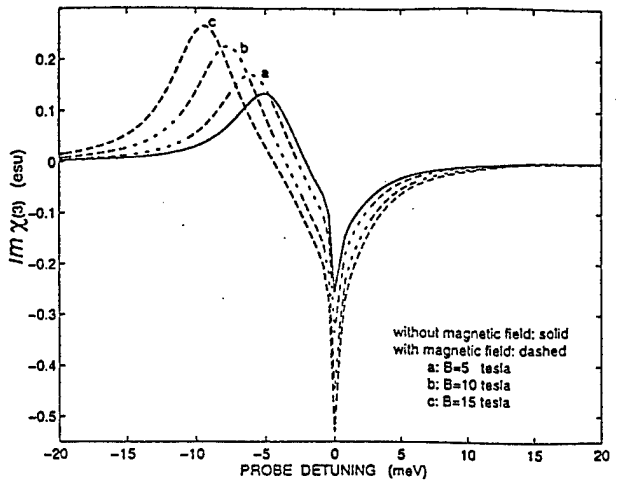



FIG. 2. The imaginary part of the third-order susceptibility as a function of the probe detuning energy for a dual-beam pump-probe experiment. The pump is set at exciton resonance for each value of magnetic field, and the longitudinal broadening parameter is assumed to be one-tenth the value of the transverse broadening parameters.

Figure 4 illustrates the dependence of third-order susceptibility on the longitudinal broadening parameters $\gamma$ (attention should be paid to the change of scale along the vertical axis compared to Fig. 2). The difference between Figs. 2 and 4 is that, in the former case, the longitudinal broadening parameter is one-tenth that of the transverse broadening parameter, whereas in the latter figure they are equal. $\mathrm{Im}\chi^{(3)}$ is extremely sensitive to the magnitude of $\gamma$: varying this damping parameter from $\gamma=0.1\Gamma$ to $\gamma=\Gamma$ changes the value of $\mathrm{Im}\chi^{(3)}$ by more than an order of magnitude.

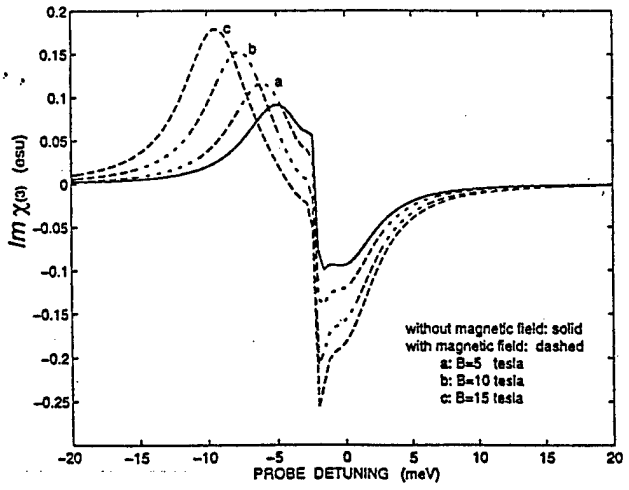



FIG. 3. The imaginary part of the third-order optical susceptibility as a function of the probe detuning energy for a dual-beam pump-probe experiment. The pump is detuned slightly below the exciton resonance for each value of magnetic field. Again, the longitudinal broadening parameter is one-tenth the value of the transverse broadening parameters as in Fig. 2.

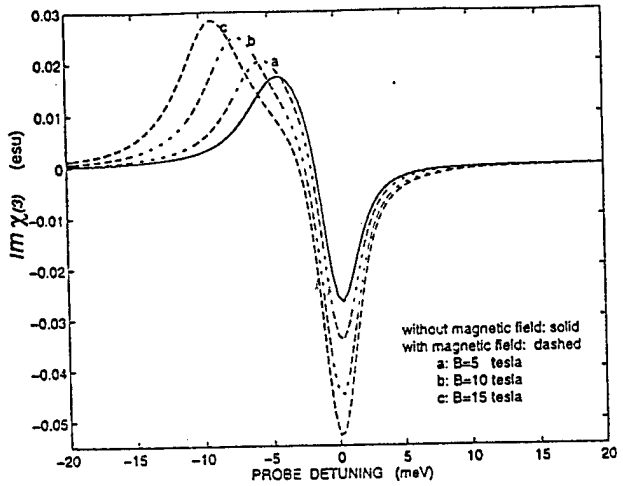



FIG. 4. The imaginary part of the third-order optical susceptibility as a function of the probe detuning energy for a dual-beam pump-probe experiment. The pump is set at exciton resonance at each value of a magnetic field as in Fig. 2, but the longitudinal broadening parameter is equal to the transverse broadening parameters. Attention should be paid to the change of scale along the vertical axis compared to Fig. 2.

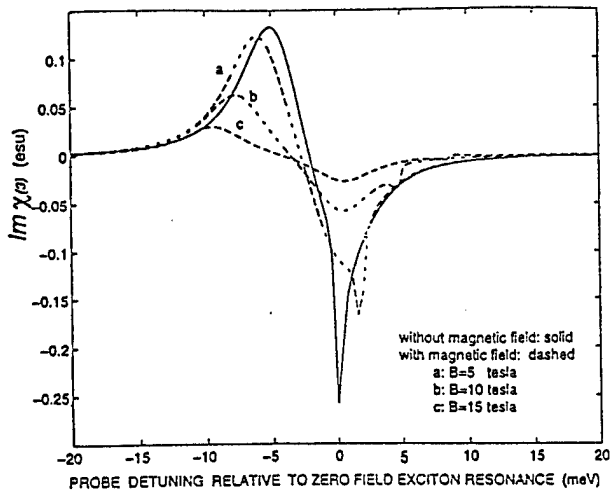



FIG. 5. The imaginary part of the third-order susceptibility as a function of the probe detuning energy for a dual-beam pump-probe experiment. The pump is now set at the exciton resonance at zero magnetic field and not retuned every time the magnetic field changes. Again, the longitudinal broadening parameter is one-tenth the value of the transverse broadening parameters as in Fig. 2. Due to the pump detuning at nonzero magnetic field, both exciton and biexciton resonances are quenched.

A pronounced negative peak is present in all of the spectra. It represents a strong transmission which is due to a saturation (or bleaching) of the excitonic state. Physically, the initial exciton population created by the pump beam tends to amplify the probe beam when its energy is tuned at or near the exciton ground state (this corresponds to the linear gain peak). A magnetic field makes the peak deeper, without significant broadening, thus enhancing transmission further. At first glance, it may surprise the reader that the peak is not shifted in frequency by the magnetic field even though the exciton binding energy depends on the magnetic field. The reason for this is that the probe beam is retuned to the exciton resonance for each value of the magnetic field, so that no frequency shift should arise.

Another feature of interest in all of these plots is in the region of positive $\mathrm{Im}\chi^{(3)}$ that corresponds to optical absorption. This absorption may be attributed to the formation of the excitonic molecule (biexciton).[4,13,14] The initial exciton population enables the probe to be more strongly absorbed when its energy matches the exciton-biexciton transition energy $\hbar(\omega_{g0}-\omega_b)$. Consequently, at zero magnetic field, the positive peak is separated from the exciton resonance by approximately $-5$ meV, which corresponds to the biexciton binding energy for this case. At a magnetic flux density of 5 T the peak separation is about $-7.5$ meV, which again corresponds to the biexciton binding energy, this time for a flux density of 5 T. Such a dependence of the energy difference between transmission (negative peak) and absorption (positive peak) on a magnetic field can be used to modulate the optical properties of a quantum wire with an external field. It can also be used as a means to determine the particular mechanism causing nonlinearity in a quantum wire. Note that the energy separation between the peaks is not seriously affected by the increasing damping (see Fig. 4) or by the detuning of the pump (see Fig. 3). As a result, this technique of modulation with a magnetic field cannot only be used to extract the mechanism responsible for nonlinearity, but also to measure biexciton binding energies and their dependences on a magnetic field.

Figure 5 shows $\mathrm{Im}\chi^{(3)}$ for four values of a magnetic field and a small longitudinal damping $\gamma=0.1\Gamma$. The difference between this case and the one presented in Fig. 2 is that now the pump beam is permanently tuned to the exciton resonance at zero magnetic field, and not retuned every time the magnetic field changes. Since the ground-state exciton energy is a function of magnetic field, the negative peak in Fig. 5 is now shifted by the applied magnetic field. Another im-

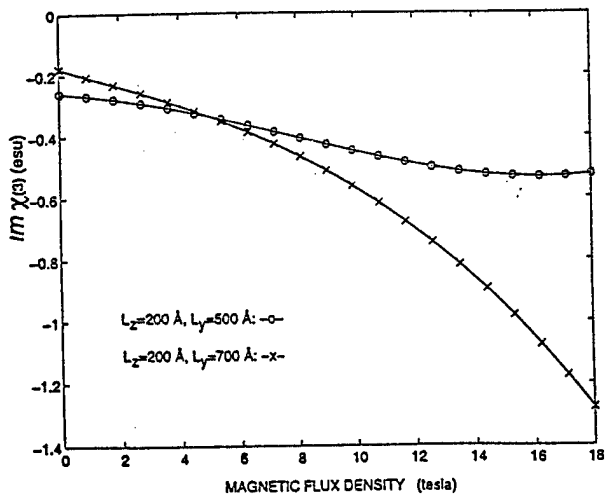



FIG. 6. The imaginary part of the third-order susceptibility as a function of an applied magnetic field for two different wire widths. The thickness of the wire is 200 Å . The effect of a magnetic field is more pronounced for a wider wire since the wave functions of the electrons and holes are "softer" and more "squeezable" by a magnetic field if the wire is wider.

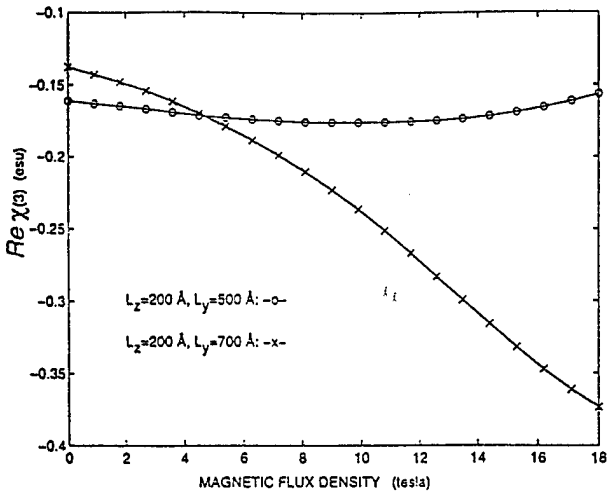



FIG. 7. The real part of the third-order susceptibility as a function of an applied magnetic field for two different wire widths. The thickness of the wire is 200 Å .

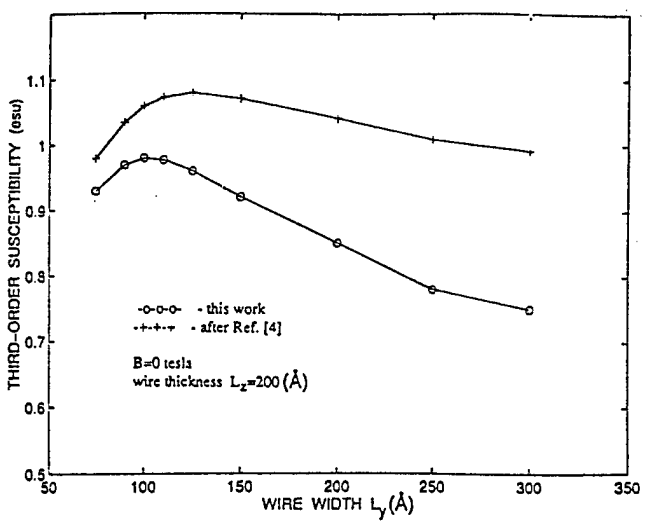



FIG. 8. Comparison of the wire-width dependence of the absolute value of the third-order susceptibility (resonant excitation) with the result given in Ref. 4. The discrepancy (15–20 %) is an aftermath of different values of material parameters used in the two studies. The thickness of the wire is 200 Å in both cases.

portant feature to notice is that when the pump is *not tuned* to an exciton resonance at a particular magnetic field, the magnitudes of both the positive and negative peaks are reduced by the magnetic field, leading to a quenching of both absorption and transmission. This effect is opposite to what is observed when the pump is tuned to the exciton resonance. However, this effect has immediate device applications such as the quantum-confined Lorentz effect.[15]

In Figs. 6 and 7 we present the imaginary and real parts of the third-order susceptibility as functions of an applied magnetic field for two different wire widths. The effect of a magnetic field is quite strong. At a magnetic flux density of 10 T, $|\mathrm{Im}\chi^{(3)}|$ is approximately three times larger than at zero field for the 700 Å -wide wire. Again, the effect of a magnetic field is more pronounced for wider wires, since in wider wires the magnetostatic localization is stronger.[11]

We have compared our results for zero magnetic field with those given in Ref. 4. The ground-state binding energies for both excitons and biexcitons are in excellent qualitative agreement. Some discrepancy can be attributed to different values of electron and hole effective masses used in our calculations and the calculations of Ref. 4. We also compare our binding energy results with the experimental observations of Refs. 16–18. Since Ref. 16 employed *T*-shaped edge quantum wires whose geometries are very different from ours, a direct quantitative comparison is not possible. Nonetheless, we find that our numerical results are within the same order of magnitude as theirs, and that their data are in excellent qualitative agreement with ours. The $\mathrm{Im}\chi^{(3)}$ curves for zero magnetic field are also consistent with those given in Refs. 4, 14, and 19. In Fig. 8, we present a direct comparison of the wire-width dependence of $\chi^{(3)}$ obtained with zero magnetic field with the result given in Ref. 4. The slight discrepancy of 15–20 % is a result of using different electron and hole effective masses in the calculation of exciton and biexciton binding energies. Other factors, such as the exact choice of the variational wave function, also contribute to the difference.

Again, in most cases, a complete quantitative comparison of the data is not possible because of the different geometry of the wires used. In our model calculations, we utilized wire sizes that are typical for structures delineated by lithography, and that correspond to the regime of moderate quantum confinement.

## IV. CONCLUSION

We have investigated the effects of a magnetic field on the third-order nonlinear susceptibility in quantum wires. The magnetic field modulates the frequency shift between the transmission peak associated with the bleaching of excitonic transitions and the absorption peak associated with the formation of excitonic molecules (biexcitons). Additionally, the field also affects the magnitudes of the peaks. These effects can be utilized for magneto-optical devices, and can also be used as a tool to probe the precise mechanism responsible for optical nonlinearity in quantum wires.

## ACKNOWLEDGMENTS

The authors are indebted to Dr. F. Madarasz and A. Ivanov for insightful discussions. They would like to thank the University of Nebraska for its hospitality. This work was partially supported by the Army Research Office under Grant No. DAAH04-95-1-0586.

*On leave from the University of Notre Dame, Notre Dame, IN 46556.

[1]F.L. Madarasz, F. Szmulowicz, F.K. Hopkins, and D.L. Dorsey, J. Appl. Phys. 75, 639 (1994).

[2]T. Ishihara, Phys. Status Solidi 159, 371 (1990).

[3]T. Ishihara and T. Goto, J. Phys. Soc. Jpn. 57, 2191 (1988).

[4]F.L. Madarasz, F. Szmulowicz, F.K. Hopkins, and D.L. Dorsey, Phys. Rev. B 49, 13 528 (1994); R.O. Klepfer, F.L. Madarasz, and F. Szmulowicz, *ibid.* 51, 4633 (1995); F.L. Madarasz, F. Szmulowicz, F.K. Hopkins, and D.L. Dorsey, *ibid.* 52, 8964 (1995).

[5]E. Hanamura, in *Optical Optics of Organics and Semiconductors*, edited by T. Kobayashi (Springer-Verlag, Berlin, 1989), pp. 2–14.

[6]A. Balandin and S. Bandyopadhyay, Phys. Rev. B 52, 8312 (1995).

[7]A. Balandin and S. Bandyopadhyay (unpublished); in Quantum Confinement II: Physics and Applications, edited by S. Bandyopadhyay *et al.* (The Electrochemical Society, Pennington, NJ, in press).

[8]N. Telang and S. Bandyopadhyay, Phys. Rev. B 48, 18 002 (1993); Appl. Phys. Lett. 62, 3161 (1993).

[9]H. Qiang, F. Pollak, C. Torres, W. Leitch, A. Kean, M. Stroscio, G. Iafrate, and K. Kim, Appl. Phys. Lett. 61, 1411 (1992).

[10]H. Ishihara and K. Cho, Phys. Rev. B 42, 1724 (1990).

[11]A. Balandin and S. Bandyopadhyay, Superlatt. Microstruct. 19, 97 (1996); A. Balandin and S. Bandyopadhyay (unpublished). For the influence of a magnetic field on electron and hole wave functions, see also Ref. 15; S. Chaudhuri and S. Bandyopadhyay, J. Appl. Phys. 71, 3027 (1992).

[12]V. Dneprovsky, N. Gushina, O. Pavlov, V. Poborchii, I. Salamatina, and E. Zhukov, Phys. Lett. A 204, 59 (1995).

[13]L. Belleguie and L. Banyai, Phys. Rev. B 47, 4498 (1993).

[14]L. Banyai, Y. Z. Hu, M. Lindberg, and S. Koch, Phys. Rev. B 38, 8142 (1988).

[15]A. Balandin and S. Bandyopadhyay, J. Appl. Phys. 77, 5924 (1995).

[16]T. Someya, H. Akiyama, and H. Sakaki, Phys. Rev. Lett. 74, 3664 (1995).

[17]R. Rinaldi, R. Cingolani, M. Lepore, M. Ferrara, I.M. Catalano, F. Rossi, L. Rota, E. Molinari, P. Lugli, U. Marti, D. Martin, F. Morier-Gemoud, P. Ruterana, and F.K. Reinhart, Phys. Rev. Lett. 73, 2899 (1994).

[18]P. Vicente, A.V. Kavokin, A. Raymond, S.G. Lyapin, K. Zekentes, D. Dur, and W. Knap, J. Phys. (France) IV 3, 323 (1993).

[19]L. Banyai, I. Galbraith, and H. Haug, Phys. Rev. B 38, 3931 (1988).

Abstract Submitted
for the MAR97 Meeting of
The American Physical Society

Sorting Category: 17.b

**Giant Dipole Effect and Second Harmonic Generation in Magnetic-Field-Bised Semiconductor Quantum Wires**

ALEXEI SVIZHENKO, *Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN.*, ALEXANDER BALANDIN, SUPRIYO BANDYOPADHYAY, *Department of Electrical Engineering, University of Nebraska, Lincoln, NE.* — We have theoretically studied the giant dipole effect in magnetic-field-biased semiconductor quantum wires. The dipoles are associated with transitions between magneto-electric subbands within the conduction band; some of these transitions are forbidden in the absence of the magnetic field. The possibility of second harmonic generation in a quantum wire biased with a magnetic field has also been examined. We will show that the simultaneous presence of a symmetric confinement (electrostatic) potential and an external magnetic field may lead to uneven charge distribution along the width of the wire which breaks the inversion symmetry in a generic quantum well or wire. This results in a non-zero third-order susceptibility and associated optical nonlinearity. This work was supported by the US Army Research Office under contract DAAH04-95-1-0586 and DAAH04-95-1-0527.

Alexander Balandin
abalandi@engrs.unl.edu
Department of Electrical Engineering
University of Nebraska, Lincoln, NE.

[X] Prefer Oral Session
[ ] Prefer Poster Session

Date submitted: December 3, 1996

Electronic form version 1.2